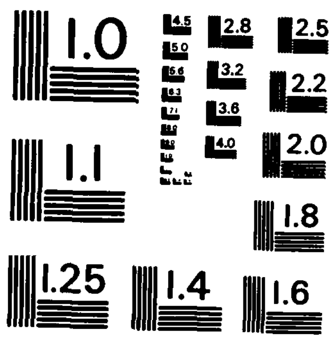
AD-A160 055 MODIFIED CONFIDENCE INTERVALS FOR THE MEAN OF AN AUTOREGRESSIVE PROCESS(U) STANFORD UNIV CA DEPT OF OPERATIONS RESEARCH B D TITUS AUG 85 TR-9 1/2

UNCLASSIFIED ARO-20927.9-MA DAAG29-84-K-0030 F/G 12/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

UNCLA[SSIFIED]
SECURITY CLASS[IFICATION OF THIS PAGE]

AD-A160 055

		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUM[BER] ARO 20927.9-MA	2. GOVT ACCE[SSION NO.] N/A	3. RECIPIENT'S CATALOG NUMBER N/A
4. TITLE (and Subtitle) MODIFIED CONFIDENCE INTERVALS FOR THE MEAN OF AN AUTOREGRESSIVE PROCESS		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Birney D. Titus		8. CONTRACT OR GRANT NUMBER(s) DAAG29-84-K-0030
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Operations Research Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE AUGUST 1985
		13. NUMBER OF PAGES 102
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

DTIC ELECTE OCT 10 1985 S D A

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

NA

18. SUPPLEMENTARY NOTES

The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Autoregressive processes, Confidence interval corrections, Cornish-Fisher expansions, Edgeworth expansions, Simulation output analysis, Time series analysis.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

PLEASE SEE NEXT PAGE

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

MODIFIED CONFIDENCE INTERVALS FOR THE
MEAN OF AN AUTOREGRESSIVE PROCESS

by

Birney D. Titus

ABSTRACT

Given $X_1, X_2, \ldots$, an asymptotically stationary autoregressive process of finite order, we wish to form a confidence interval for the steady state mean. The usual confidence intervals have one sided coverage probability errors which are $O(n^{-1/2})$, where n is the sample size. We derive a first order correction which, under reasonable assumptions, reduces the one sided errors to $O(n^{-1/2})$ and a second order correction which reduces the errors to $O(n^{-1})$. These corrections produce confidence intervals with more accurate coverage. The motivation for and a principal application for these results is the analysis of simulation output data.

Keywords: Autoregressive processes
Confidence interval corrections
Cornish-Fisher expansions
Edgeworth expansions
Simulation output analysis
Time series analysis

ARO 20927.9-MA

MODIFIED CONFIDENCE INTERVALS FOR THE
MEAN OF AN AUTOREGRESSIVE PROCESS

by

Birney D. Titus

TECHNICAL REPORT NO. 9

August 1985

Prepared under the Auspices

of

U.S. Army Research Contract*
DAAG29-84-K-0030

DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

*This research was also partially supported under
National Science Foundation Grant MCS-8203483.

85 10 9 205

Contents

CONTENTS

1

Introduction

Statement of the problem

~~Our~~ The author's motivation is to find asymptotically more accurate confidence intervals for the steady state mean of a simulated process. By this ~~we~~ he means that the coverage probability error for the confidence intervals we derive should be of lower order than that of standard confidence intervals. There are several standard methods of setting confidence intervals in simulations, including the regenerative method, batch means, and time series methods. ~~We will~~ He focuses on improved confidence intervals for the mean of an autoregressive process, and as such our results are useful outside of a simulation setting.

Improved methodology for setting simulation confidence intervals is an area of active research. A recent survey article, LAW AND KELTON (1984), states

One of the most important but difficult problems encoun-

Additional keywords: Time series analysis; Cornish-Fisher expansions; Edgeworth expansions.

tered in a real-world simulation study is that of constructing a confidence interval (c.i.) for the steady state mean μ of a stochastic proces. The information contained in such a c.i. provides the decision maker with a measure of how precisely μ is known. Constructing the c.i., however, is difficult because the output data from a simulation are in general non-stationary and autocorrelated, so that *direct* application of the techinques of classical statistics is precluded.

Almost none of the work on confidence intervals aims to improve the asymptotic *order* of confidence interval accuracy, however. If we use standard methods to set a nominal 90 percent confidence interval for the mean of the process, the true proportion of the time that the actual mean falls within our interval will be 0.9 plus an error term which is typically $O(n^{-1})$. The error in half-coverage, that is the difference between 0.45 and the true proportion of the time that the actual mean falls in one half of the nominal interval, is of order $O(n^{-1/2})$. Most work attempts to improve the constant factor implicit in this $O(n^{-1/2})$. We will derive a first order correction which, under reasonable assumptions, reduces the one sided errors to $o(n^{-1/2})$ and a second order correction which reduces these errors to $o(n^{-1})$.

Importance of the problem

In the usual statistical context one feels that a sample size of $n = 30$ is reasonably large—certainly the t-distribution with 29 degrees of freedom closely approximates the normal distribution, and one regards normal theory as adequate for forming confidence intervals for the mean. In forming a confidence interval for the mean of an asymptotically stationary process, however, one may need $n = 10{,}000$ or greater. To illustrate this, a Monte Carlo study (discussed in Chapter 5) on the expected stationary waiting time in an M/M/1 queue in light traffic (traffic intensity of 0.5) showed among other things that the true probability that the sample mean was greater than the upper limit of a 90 percent confidence interval was not 0.05 but about 0.14. Various criteria have been suggested for evaluating confidence intervals, but reasonable accuracy is the *sine qua non* of confidence intervals.

For this reason, it is desirable to find more accurate methods to form confidence intervals for small samples. As mentioned above, more accurate *half* coverage is desirable. For example, we might wish to perform a one-sided test on the estimated mean.

Basic tools and assumptions

We assume we are given a process $x = \{x_i : i \geq 1\}$ which satisfies a stable autoregressive difference equation,

$$a_0(x_i - \mu) + \cdots + a_k(x_{i-k} - \mu) = \epsilon_i \tag{1}$$

for $i = k+1, \ldots, n$. The ϵ's of the sequence $\{\epsilon_i : i \geq k+1\}$ are zero-mean, independent and identically distributed random variables with moments of all orders and a density (with respect to Lebesgue measure) which is positive on an interval. We take the point of view, then, that the simulator or statistician has determined that such a model is appropriate. If the prediction errors of the model (epsilons) may only be assumed uncorrelated and not independent, the first order correction still applies with minor modifications. If the model of actual interest is in continuous time, this of course necessitates some method like sampling at discrete intervals. In the regenerative method of forming confidence intervals for the steady state value of $Ef\,(x(t))$, where $x(t)$ is a regenerative process in continuous time with regeneration times $\{t_i : i \geq 0\}$, one would proceed as follows. Let τ_i be defined as the regenerative cycle length,

$$\tau_i \triangleq t_i - t_{i-1},$$

and let y_i be defined by

$$y_i \triangleq \int_{t_{i-1}}^{t_i} f\,(x(t))\, dt.$$

The mean we want to estimate is $Ey_i/E\tau_i$, and we may use the Delta method to derive the variance of the natural estimate $\bar{y}_n/\bar{\tau}_n$, where $\bar{y}_n$ and $\bar{\tau}_n$ are the respective sample means. (See BILLINGSLEY (1979) for the Delta method). If we are not using the regenerative method, we could sample the process at fixed intervals, which is to say that τ_i is constant, but it may be more convenient to sample the process at random intervals. For appropriate choices of the random times $\{t_i : i \geq 0\}$ the joint sequence $\{(y_i, \tau_i) : i \geq 0\}$ as defined above will still be asymptotically stationary and will still obey the bivariate Central Limit Theorem, even though it is not an independent sequence. In this case, $Ey_i/E\tau_i$ is still the natural point estimate, and the Delta method may be used exactly as before to derive the same variance constant as one would derive using the regenerative method. The only difference is that the covariance matrix of the process $\{(y_i, \tau_i) : i \geq 0\}$ is not as easy to estimate. One may, however, estimate this covariance matrix using a bivariate autoregressive model, as in JOW (1983), or some other method. If a bivariate autoregressive method is used, then the natural extensions of the methods of this thesis may be used to obtain more accurate confidence intervals, though this extension to the multivariate autoregressive method is not elaborated here. See FOX AND GLYNN (1983) for further discussion.

The constant μ of (1) is the asymptotic mean of the x process, and a_0 is assumed equal to one. Note that the sequence of interest, $x = \{x_i : i \geq 1\}$, need not be stationary, but it will be asymptotically stationary. The requirement that the difference equation be strictly stable is equivalent to the technical condition that all roots of the characteristic polynomial

$$a_0 z^k + \cdots + a_k z_0$$

lie strictly inside the unit circle (see PRIESTLEY (1981)).

If the usual pivot or test statistic based on n sample points is t_n, then the first order pivot we will derive will have the form $T_1 = t_n + \theta_n n^{-1/2} + \rho_n t_n^2 n^{-1/2}$, and the second order pivot will be

$T_2 = T_1 + \nu_n t_n n^{-1} + \omega_n t_n^3 n^{-1}$. The pivots T_1 and T_2 depend on n, but we suppress this dependence in the notation. T_1 differs from a standard normal random variable by an error term which is "little oh" of $n^{-1/2}$ in probability, written $o_p(n^{-1/2})$, and T_2 by a term which is $o_p(n^{-1})$. Recall that $y_n = o_p(n^{-s})$ means that $n^s y_n$ goes to zero in probability. These test statistics will be the basis of confidence intervals with coverage error probability of asymptotically lower order.

The basic tools behind our derivations are the Edgeworth expansion, and the Cornish-Fisher expansion. KENDALL AND STUART (1977) contains a good introduction to these concepts. The Edgeworth expansion is an asymptotic expansion for the central limit theorem. If $\Phi(z)$ denotes the standard normal distribution function and $\phi(z)$ the standard normal density, the most basic form of Edgeworth expansion is

$$P\{n^{-1/2}\textstyle\sum_1^n (y_i - Ey)/\sigma \le z\} = \Phi(z) + \phi(z)\left[-\kappa_1 - \kappa_3(z^2-1)/6 + \cdots\right],$$

where $\{y_i : i \ge 1\}$ is an independent and identically distributed sequence satisfying certain moment and smoothness assumptions. The two κ's shown above are $O(n^{-1/2})$. Less basic forms of the expansion will allow the y's to be non-independent and allow the probability that the normalized sample mean lies in more arbitrary Borel sets to be estimated. In addition, we usually are concerned not with the normalized sample mean but with a function of several normalized sample means.

The Cornish-Fisher expansion amounts to a polynomial transformation, which transforms the asymptotically normal statistic t_n to a statistic with a distribution closer to that of the standard normal distribution. In fact, the first order corrected statistic T_1 above is a Cornish-Fisher expansion of t_n, and T_2 is a Cornish-Fisher expansion of T_1.

Previous work

There are several large areas of research which provide a basis for this thesis or which are related to it.

The first such area is that of Edgeworth expansions. Edgeworth's paper appeared in 1904. The book BHATTACHARYA AND RAO (1976) provides very good background for the independent and identically distributed case, as does the paper BHATTACHARYA AND GHOSH (1978). TANIGUCHI (1984) deals with the time series context, and GOTZE AND HIPP (1983) prove the validity of Edgeworth expansions for quite general functions of weakly dependent random variables.

The original paper of Cornish and Fisher appeared in 1927. See HILL AND DAVIS (1968) for a more recent study.

Another major area of research is confidence interval methodology in general. FISHMAN (1978) provides background and presents the autorregressive method of simulation output analysis. JOW (1982) details the autoregressive method for vector processes, and the articles LAW AND KELTON (1982, 1984) survey simulation oriented work, and contain further references. There is also more statistically oriented research, such as that of Efron cited below. The electrical engineering literature contains work relevant to confidence intervals, too. For example, THOMSON (1982) discusses sophisticated methods for spectrum estimation, and though confidence intervals are not specifically mentioned, the problem of estimating the variance of the sample mean of an asymptotically stationary process may be viewed as one of spectrum estimation–see JOW (1983) and PRIESTLEY (1981).

Finally, there is other work which deals with improving the *order* of confidence interval accuracy. JOHNSON (1978) is perhaps the first such article. He derives a first order correction for the usual t-statistic in an independent and identically distributed context. GLYNN (1982a) extends this to a second order correction for ratio estimation, and the methods of EFRON (1984b) center around the parametic bootstrap. See also ABRAMOVITCH AND SINGH (1985) and the articles cited there.

Overview

In Chapter 2 we will develop machinery to help us derive the actual corrections, though the results of this chapter apply quite generally to asymptotically stationary sequences. To use the Cornish-Fisher expansions, we need to estimate moments and cumulants of some function $f(x)$, where x is a vector of sample means of an asymptotically stationary sequence. After expanding f in a Taylor series, we thus need to estimate various joint moments and cumulants of sample means. One aspect of this is the need to sort out the contributions to a given cumulant in terms of (negative) powers of the sample size, n. It is primarily the second section of Chapter 2, "Moment identities," which details this machinery. The first section, "Cumulant and moment bounds," facilitates the derivations of the second section.

In this first section, a result of JAMES (1955, 1958) and JAMES AND MAYNE (1962) is simplified and extended to weakly dependent random variables. The result of James asserts that the jth cumulant of a polynomial in the sample average $a_0 + a_1(n^{-1}\sum_1^n x_i)^1 + \cdots + a_k(n^{-1}\sum_1^n x_i)^k$ for independent and identically distributed x_i is $O(n^{1-j})$ as n becomes infinite, and this result is referred to by BHATTACHARYA AND GHOSH (1978) as "an important combinatorial result."

Chapter 3 shows how to compute certain infinite sums associated with sequences (on $\mathbf{Z}_+^2, \mathbf{Z}^2$, etc., where $\mathbf{Z}$ is the set of integers and $\mathbf{Z}_+$ is the set of positive integers) which obey homogeneous or non-homogeneous difference equations. Many expressions for cumulants of functions of the mean of an autoregressive process are of this type.

Chapter 4 gives the actual derivation of the corrections, along with an algorithmic summary. Given the methods of Chapters 2 and 3 the derivation is relatively mechanical, though by no means completely so. Making the Cornish-Fisher correction entails first computing the usual test statistic, t_n. One must also estimate the third and fourth moments of the residuals ϵ_i in the autoregressive model. If n is the sample size, estimation of covariances of the underlying sequence and of the moments of ϵ_i requires work proportional to $O(n)$. The rest of the computation requires work of order $O(k^3)$, where k is the order of the autoregressive model. The first order correction is fairly simple, and despite the number of steps the second order correction is not too computer-intensive.

Finally, we present numerical results in Chapter 5. We consider two genuine autoregressive processes with independent and identically distributed residuals. In one case, the residuals have a smooth density. The second example is the same, except that the residuals have a lattice distribution. The third model is the M/M/1 queue, with is not a finite order autoregressive process. The fourth model is the discrete analog of the M/M/1 queue, that is, a random walk on the nonnegative integers. This variety of models is included to test, at least in a few instances, the sufficiency of our sufficient conditions and the necessity of our necessary conditions.

2

The algebra of moments

In this chapter we will develop tools which are very useful for simplifying moment expressions. For example, to calculate the usual test statistic (or "pivot") we need to estimate both the mean of the process of interest and covariances. *If the estimate of the mean is* $\bar{x}_n$ and that of the covariance at lag 0 is $\hat{R}(0)$, in the first order correction we will need to estimate the covariance of $\bar{x}_n$ and $\hat{R}(0)$. This sort of covariance or mixed moment is what we mean by "moment expression."

First we will obtain a result which shows, under appropriate conditions, that the Taylor expansion of a function of a sample mean has cumulants of the same order as the corresponding cumulants of a sample mean. This result and its generalizations will help us to derive the higher order moment identities in the second part of this chapter.

Cumulant and moment bounds

2.1 Introduction. Let S_n denote the sum $\sum_1^n x_i$, where the x_i are mixing and asymptotically stationary (in a sense to be detailed later) with asymptotic mean zero. If we wish to approximate the cumulants of $\sqrt{n}(f(S_n/n)-f(0))$, we are lead to consider the cumulants of a Taylor expansion: $\alpha_1 S_n n^{-1/2}+\cdots+\alpha_k S_n^k n^{1/2-k} = p(S_n)$. In case the sequence $\{x_i : i \geq 1\}$ is independent and identically distributed and $p(S_n) = S_n n^{-1/2}$, then the cumulant generating functions of the polynomial p and of x are related by the formula $\psi_p(\varsigma) = n\psi_x(\varsigma n^{-1/2})$, which shows that the jth cumulant for the normalized polynomial p, $\kappa_{j,n}$, is $O(n^{1-j/2})$. JAMES (1955, 1958) and JAMES AND MAYNE (1962) have shown using the machinery of Fisher's k-statistics that the same result holds for general $p(S_n)$ when the the sequence $\{x_i : i \geq 1\}$ is independent and identically distributed.

Our purpose in this section is to give a simple proof of this fact and some of its extensions.

Instead of considering the polynomial p, we take $\sqrt{n}p$, that is $\alpha_1 S_n + \cdots + \alpha_k S_n^k n^{1-k}$, and we will show that the cumulants are $O(n)$. We will begin with the case in which the x's are independent, but we do not require zero means or even identical distributions.

2.2 Review of cumulants. We record here several facts about cumulants.

(1) The joint cumulant $\text{cum}(y_1, \ldots, y_k)$ is given by

$$\sum_{\nu} (p-1)!(-1)^{p-1} \mu_{\nu_1} \cdots \mu_{\nu_p}$$

where ν denotes a partition of $\{y_1, \ldots, y_k\}$ into p sets, where μ_{ν_i} is the expected product of those y's in the ith partition, and where the sum is taken over all partitions ($p = 1, \ldots, k$). For example, $\text{cum}(y_1, y_2) = Ey_1 y_2 - Ey_1 Ey_2$.

(2) By definition, $\text{cum}(y_1, \ldots, y_k)$ is the coefficient of $s_1 \cdots s_k$ in the Taylor expansion of $\log E \exp[i(s_1 y_1 + \cdots + s_k y_k)] = \psi_{\mathbf{y}}(\mathbf{s})$ The function ψ is the ***cumulant generating function.*** The jth cumulant of the univariate distribution of y is the jth derivative the cumulant generating function at zero, $D^j \psi_y(0)$.

(3) The jth cumulant of y, which is denoted by κ_j or $\kappa_j(y)$, is equal to $\text{cum}(y, \ldots, y)$ (j times). This is evident from the fact that $\kappa_j = D^j \psi_y(0) = D^{1,\ldots,j} \psi_{\mathbf{y}}(\mathbf{0}) = \text{cum}(y, \ldots, y)$.

(4) The cumulant $\text{cum}(xy_1, y_2, \ldots, y_n) = 0$ if x is independent of the y's and $Ex = 0$, because Ex then factors out of all the moments in the expansion of (1) above.

(5) $\text{cum}(y_1, \ldots, y_j) = 0$ whenever the y's may be partitioned into two groups independent of each other. In this case, supposing $(y_1, \ldots, y_l)$ is independent of $(y_{l+1}, \ldots, y_j)$,

$$\log E \exp(i\mathbf{s}^T \mathbf{y}) = \log E \exp(s_1 y_1 + \cdots + s_l y_l) + \log E \exp(s_{l+1} y_{l+1} + \cdots + s_j y_j),$$

whose Taylor expansion has no $s_1 \cdots s_j$ term.

(6) The jth cumulant of $\alpha_1 y_1 + \cdots + \alpha_l y_l = \alpha^T \mathbf{y}$ is a linear combination of jth order cumulants which we could write as

$$\sum_{|\nu|=j} \beta_\nu \text{cum}(y_\nu),$$

where ν is a multiindex, $\nu = (\nu_1, \ldots, \nu_j)$, $|\nu| = j$ means that ν has j ordered elements, and $\text{cum}(y_\nu) = \text{cum}(y_{\nu_1}, \ldots, y_{\nu_j})$. If ψ_1 is the cumulant generating function of $\alpha^T \mathbf{y}$ and ψ_2 that of $\mathbf{y}$, then $\psi_1(s) = \psi_2(\alpha s)$. Taking jth derivatives on both sides with respect to s now proves the assertion.

(7) The cumulant is symmetric and multilinear. For example, if $S = \sum_1^n x_i$,

$$\begin{aligned} \text{cum}(S, S^2/n, S^3/n^3) &= n^{-3}\text{cum}(S, S^2, S^3) \\ &= n^{-3} \sum \text{cum}(x_{i_1}, x_{i_2}x_{i_3}, x_{i_4}x_{i_5}x_{i_6}). \end{aligned}$$

We are now ready to prove the theorem for the independent case.

2.3 Theorem. *Let $\{x_i : i \geq 1\}$ be an independent sequence and let j and k be positive integers. Suppose that the set of moments $Ex_{i_1} \cdots x_{i_l}$ is bounded for $l \leq jk$. Then the jth cumulant of $\alpha_1 S + \cdots + \alpha_k S n^{1-k}$, where $S = \sum_1^n x_i$, is $O(n)$ as $n \to \infty$.*

Proof. From fact (6) above, the jth cumulant of $\alpha_1 S + \cdots + \alpha_k S n^{1-k}$ is a linear combination $\sum_{|\nu|=j} \beta_\nu \text{cum}(y_\nu)$, where $y = (S, \ldots, S^k n^{1-k})$ and where the coefficients β_ν do not depend on n. It therefore suffices to show that each $\mathrm{cum}(y_\nu)$ is $O(n)$. The expression $\mathrm{cum}(y_\nu)$ means $\mathrm{cum}\left(S^{\nu_1}n^{1-\nu_1},\ldots,S^{\nu_j}n^{1-\nu_j}\right)$ if $\nu=(\nu_1,\ldots,\nu_j)$. Notation may obscure the simple idea here, so let us suppose we are dealing with $\mathrm{cum}\left(S,S^2/n,S^2/n,S^3/n^2\right)$. That the same argument applies in general will be obvious.

Fact (7) of the previous subsection shows that this cumulant is

$$n^{-4}\sum \mathrm{cum}(x_{i_1},x_{i_2}x_{i_3},x_{i_4}x_{i_5},x_{i_6}x_{i_7}x_{i_8}).$$

The indexes $(i_1,\ldots,i_8)$ come from the set of eight-tuples of positive integers less than or equal to n, but fact (5) shows that this cumulant is often zero. In fact, if the four arguments split into two independent groups, the cumulant is zero. Therefore we must specify three constraints (the number of arguments minus one) to make it nonzero. That is, if each argument is a node and if we regard two nodes as connected whenever the arguments are dependent, we must have a connected graph. Two arguments in our context are dependent only when they share an index—for example x_{i_1} and $x_{i_2}x_{i_3}$ are dependent only if $i_1=i_2$ or $i_1=i_3$. Note that the possible number of such constraints does not depend on n. Thus the cumulant is nonzero only if the indexes satisfy at least one set of three equality constraints from a fixed number of j possible sets. It follows that the number of nonzero terms in the sum $\sum \mathrm{cum}(x_{i_1},x_{i_2}x_{i_3},x_{i_4}x_{i_5},x_{i_6}x_{i_7}x_{i_8})$ is $O(n^{8-3})=O(n^5)$. Because of our boundedness assumption on moments, and taking into account the denominators in the cumulants' arguments, the total sum is $O(n^5/n^4)=O(n)$ as required.

In general, each new block $x_{j_1} \cdots x_{j_l} n^{1-l}$ which is an argument of the cumulant increases the size of the space of possible indexes by a factor of n^l, but by introducing one additional necessary constraint and dividing by n^{l-1}, the growth of the sum remains at $O(n)$. Furthermore, if there is only one block, no constraints are needed. □

Before proving the theorem for non-independent sequences, we will review the concept of ***mixing***.

2.4 Review of mixing. The useful hypothesis of mixing generalizes the idea of independence. Let $\{x_i : i \geq 1\}$ be a sequence of random variables and suppose the event A is in the sigma field $\sigma(x_1, \ldots, x_k)$, while $B \in \sigma(x_{k+n}, \ldots)$. Then if the absolute difference of probabilities$|Pr(AB) - Pr(A)Pr(B)|$ is less than or equal to α_n uniformly for all such A and B, we say that the x's are ***mixing***, provided $\alpha_n \to 0$. The α_n are referred to as ***mixing constants.***

Even to prove a central limit theorem for the x's, we would require mixing or something which implies it. See for example BILLINGSLEY (1968), pages 166 and 174. One may also show (BILLINGSLEY (1979), p. 317) that if the random variable y is measurable $\sigma(x_1, \ldots, x_k)$ while z is measurable $\sigma(x_{k+n}, \ldots)$, and if y and z have bounded fourth moments, then $|Eyz - EyEz| \leq 8(1 + Ey^4 + Ez^4)\alpha_n^{1/2}$.

2.5 Theorem. *Let $\{x_i : i \geq 1\}$ be a sequence such that (1) for some positive integers j and k, the set of moments $Ex_{i_1} \cdots x_{i_l}$*

is bounded for $l \leq 4(jk-1)$ *and (2) the sequence is mixing with mixing constants* $\alpha_n = O(n^{-2(j-1+\epsilon)})$ *for some* $\epsilon > 0$. *Then the* j*th cumulant of* $\alpha_1 S + \cdots + \alpha_k S n^{1-k}$, *where* $S = \sum_1^n x_i$, *is* $O(n)$ *as* $n \to \infty$.

Remark. In applying the lemma of BILLINGSLEY cited in the last subsection, the random variables y and z will be products of at most $kj-1$ x's, say $y = x_\nu = x_{\nu_1} \cdots x_{\nu_m}$, and similarly $z = x_\rho$. Then the lemma together with the hypotheses of our theorem guarantee $|Ex_\nu x_\rho - Ex_\nu Ex_\rho|$ is uniformly $O(g^{-(j-1+\epsilon)})$ whenever the minimum ρ_i exceeds the maximum ν_i by at least g.

Proof. As in Theorem 2.3 we only need to show that each $\mathrm{cum}(y_\nu)$ is $O(n)$, and this reduces to considering the same sort of sum of cumulants. The only difference is that terms like

$$\sum \mathrm{cum}(x_{i_1}, x_{i_2}x_{i_3}, x_{i_4}x_{i_5}, x_{i_6}x_{i_7}x_{i_8})$$

may never be zero, but with a strong enough mixing condition the sum of such terms will still be $O(n)$, as we shall see.

To be specific, suppose again that we are dealing with

$$n^{-4}\sum \mathrm{cum}(x_{i_1}, x_{i_2}x_{i_3}, x_{i_4}x_{i_5}, x_{i_6}x_{i_7}x_{i_8})$$

which we want to be $O(n)$. We have $t = 4$ arguments of the cumulant, $d = 8$ dimensions to the space of indexes, $c = t - 1 = 3$ necessary constraints for a nonzero cumulant in the independent case, and $p = 4$ total powers of n in the denominator, which have been factored out of the sum. The relation $d - c - p = 1$ always holds, and this guarantees that the sum is $O(n)$ in Theorem 2.3.

Suppose that $i_1 \leq \cdots \leq i_8 = i_d$. It is enough to show that the sum over such terms is $O(n)$, as long as we can do the same for each ordering of the i's. We will need to consider the cth smallest gap $i_k - i_{k-1}$ in the index $i = (i_1, \ldots, i_d)$, which we call g. If the gap g is "large", then x's whose indexes differ by g or more will be nearly independent. If such x's actually were independent, and if the cth smallest gap in $(i_1, \ldots, i_d)$ is g, then the corresponding cumulant would be zero. In our specific example ($c = 3$), this means that only two pairs of x's or one triple could be dependent, which in turn implies that one of the cumulant's arguments would be independent of the rest, exactly as in the proof of Theorem 2.3. With our mixing condition, the cumulant will be nearly zero. Therefore, we must answer two questions:

(1) If the cth smallest gap in $(i_1, \ldots, i_d)$ is g, then *how* small is the corresponding cumulant, and

(2) How many points $(i_1, \ldots, i_d)$ have cth smallest gap equal to g?

To answer the first question, we expand the cumulant in terms of moments, as in (2.2.1). If we factor each expected product in the moment expansion as though the x's whose indexes differ by g or more were independent, we arrive at an expression which is identically zero (because g is the cth smallest gap). But now each time we factor such an expected product we incur an error which is uniformly $O(g^{-(j-1+\epsilon)})$—see the remark preceding the proof of this theorem. Note that each term in the cumulant expansion is the product of at most d moments, and each moment is the expected product of at most d x's. If $b > 1$ is our uniform bound on x-moments of order d or less, then each term of the cumulant expansion is bounded by b^d in absolute value. There are exactly d x's in each term of the expansion (each index occurs once), and each time we factor we induce an error bounded by $O(g^{-(j-1+\epsilon)})$. The total absolute error induced by factoring expected products is therefore $O(g^{-(j-1+\epsilon)})$. This answers the first question: cumulants whose cth smallest gap is g are uniformly $O(g^{-(j-1+\epsilon)})$ in absolute value (*before* multiplying by the factor $n^{-p} = n^{-4}$ in our example).

Next we must bound the number of $(i_1, \ldots, i_d)$ with cth smallest gap equal to g. Recall that $i_1 \leq \cdots \leq i_8 = i_d$, and let $g_1 = i_1, \ldots, g_d = i_d - i_{d-1}$. The cth smallest gap is the cth smallest of $g_2, \ldots, g_d$. Again, it suffices to find a bound based on the assumption that $g_2, \ldots, g_{c+1}$ are less than or equal to $g_{c+2}, \ldots, g_d$. There are g^c ways of assigning values in $\{1, \ldots, g\}$ to the c smallest g's and $(g-1)^c$ ways of assigning the values $\{1, \ldots, g-1\}$. Therefore there are $O(g^{c-1})$ ways to assign $\{1, \ldots, g\}$ such that the cth smallest g_i is exactly equal to g. Some of these ways may not satisfy the constraint $\sum g_i = n$, but that does not affect our upper bound. There are fewer than n^{d-c} ways to assign the remaining g_i's, hence the number of $(i_1, \ldots, i_d)$ with cth smallest gap of g is $O(g^{c-1} n^{d-c})$. This answers the second question stated on above.

We know already that the number of cumulants in the sum with $g = 0$ is exactly n^{d-c}. Therefore the entire cumulant sum, after dividing by the denominator $n^p = n^{d-c-1}$ is bounded in absolute value by a constant times $n^{-(d-c-1)}(n^{d-c} + n^{d-c}\sum_{g=1}^{\infty} g^{c-1}/g^{j-1+\epsilon}) = O(n)$ if $j - 1 + \epsilon > c$. For jth order cumlulants, the number of constraints c is $j - 1$, so the cumulant sum is $O(n)$. □

2.6 Remark. Note in particular that the jth cumulant of S, $\mathrm{cum}(S, \dots, S)$, is $O(n)$, where $S = \sum_1^n x_i$, provided the mixed x-moments of order $4(j-1)$ are uniformly bounded (we assume $j > 1$).

We also have $\mathrm{cum}(S, \dots, S, x_n) = O(1)$, since the only contributing terms are those for which the indexes $i_1, \dots, i_j$ are all approximately equal to n. In general,

$$\mathrm{cum}(S, \dots, S, \textstyle\prod_{l=1}^{m} x_{n-i_l}) = O(1).$$

Here the i_l are fixed nonnegative integers, not necessarily distinct. For this to hold (assuming S repeats j times) it suffices that the mixed x-moments of order $4(j + m - 1)$ be uniformly bounded and the mixing constants α_m be of order $O(m^{-2(j+\epsilon)})$.

In this connection we have the following corollary.

2.7 Corollary. *Let $S = \sum_1^n x_i$. Then for positive integer k the moment $ES^k x_{n+i_1} \cdots x_{n+i_p}$ is of order n^p, where p is at most $[k/2]$, provided*

(1) $E|x_i| = O(n^{-2})$.

(2) The kth cumulant of S is $O(n)$.

(3) $\mathrm{cum}(S, \dots S, x_{n-i_1} \cdots x_{n-i_l}) = O(1)$ *(S repeats k or fewer times as an argument).*

(4) x has uniformly bounded mixed $l + k$-order moments.

Proof. Use induction on $k+l$. If $k+l \leq 1$, the conclusion follows immediately from (1) and (4).

For $k \geq 2$ and $l = 0$ we observe that the kth cumulant of s is $O(n)$, expand the cumulant in terms of moments, and apply the induction hypothesis. The first term in the cumulant expansion is ES^k, while the other terms are products $\prod ES^{\nu_i}$ with $\sum \nu_i = k$—so the induction hypothesis proves this case. For general $k + l \geq 2$, consider $\operatorname{cum}(S, \ldots, S, x_{n-i_1} \cdots x_{n-i_l})$, which is $O(1)$ by hypothesis. □

2.8 Remark. In a sense these results may not seem as good as one might hope, and they can in fact be strengthened. Suppose we are interested in the second cumulant of $S/n + S^2/n^2$. Then we have shown that $\operatorname{cum}(S/n, S/n)$, $\operatorname{cum}(S/n, S^2/n^2)$, and $\operatorname{cum}(S^2/n^2, S^2/n^2)$ are all $O(1/n)$. This means that the least significant term in the Taylor expansion has as much impact as the most significant term. We can sharpen the result if the sequence $\{x_i : i \geq 1\}$ has asymptotic mean zero, which will normally be true in the cases of interest to us.

2.9 Corollary. *Let $\{x_i : i \geq 1\}$ be an independent sequence of random variables with zero mean and uniformly bounded moments of order $p = \sum_1^j p_i$ (the p_i are positive integers), and let $S = \sum_1^n x_i$. Then $\operatorname{cum}(S^{p_1}, \ldots, S^{p_j})$ is $O(n^{[p/2]})$ and also $O(n^{p-j+1})$. ($[p/2]$ means the greatest integer in $p/2$.)*

Proof. The $O(n^{p-j+1})$ bound follows from the argument of Theorem

2.3: there must be $j-1$ constraints on the indexes of

$$\mathrm{cum}\Big(\prod_{j=1}^{p_1} x_{i_1}, \cdots\Big)$$

for a nonzero cumulant. It is also true, however, that each index i_m must be equal to at least one other index. If for example the index i_m appears only once, then $Ex_{i_m} = 0$ factors out of each term in the expansion of the cumulant as a sum of moments. Thus for a nonzero contribution we require $\lceil p/2 \rceil$ constraints, and the sum is therefore $O(n^{\lfloor p/2 \rfloor})$. □

Either of the two bounds in the statement of the corollary may be sharper. Our new bound shows, for independent zero mean x's, that $\mathrm{cum}(S/n, S^2/n^2)$ and $\mathrm{cum}(S^2/n^2, S^2/n^2)$ are $O(n^{-2})$, which is an improvement.

The simplest asymptotically stationary case occurs when the x's are independent with zero mean, except that $Ex_1 \neq 0$. In this case for a nonzero cumulant each index of $\mathrm{cum}\Big(\prod_{j=1}^{p_1} x_{i_1}, \cdots\Big)$ must be equal to another index or equal to one, and this still requires $\lceil p/2 \rceil$ constraints. In the dependent asymptotically stationary (zero mean) case, this means each index must be near another index or near zero.

For how many sets of indexes $(i_1, \ldots, i_p)$ is each index between one and n and also within g units of zero or of another index? We can bound this number by $cn^{\lfloor p/2 \rfloor} g^p$ where c is independent of n and g. Therefore the number with this minimum gap equal to g is bounded by $c'n^{\lfloor p/2 \rfloor} g^{p-1}$. Suppose $E|x_m| = O(m^{-l})$ and the mixing constants α_m are $O(m^{-2l})$. We then find that the cumulant of the corresponding selection of x's, namely $\operatorname{cum}\left(\prod_{j=1}^{p_1} x_{i_1}, \cdots\right)$ is $O(g^{-l})$.

Combining these observations with the argument of Theorem 2.5 gives the following.

2.10 Corollary. *Let $p = \sum_1^l p_i$, where each p_i is a positive integer, and $S = \sum_1^n x_i \quad (p \geq 2)$. Suppose that*

(1) mixed moments of x's of order $4(p-1)$ are uniformly bounded.

(2) The x's are mixing with mixing constants

$$\alpha_m = O(m^{-2(p+\epsilon)}) \quad (\epsilon > 0)$$

(3) $E|x_m| = O(m^{-(p+\epsilon)})$.

Then $\operatorname{cum}(S^{p_1}, \ldots, S^{p_l})$ *is* $O(n^{\lfloor p/2 \rfloor})$. □

Note that the argument of Theorem 2.5 implies that the cumulant above is $O(n^{p-l+1})$ if conditions (2) and (3) above are replaced by the requirement that the mixing constants $\alpha_m = O(m^{-2(l-1+\epsilon)})$.

We need to make two important remarks.

2.11 Remark. Suppose some or all of the arguments to the cumulant are multiplied by a product of x_{n-i}'s. In other words, each argument is $S^{p_j} \prod_{l=1}^{m_j} x_{n-i_{l,j}}$, where we allow the product to be empty or $p_j = 0$ (but not both). If

(1) The sum $\sum p_j$ is denoted by p,

(2) the total number of indexes is $d = p + \sum m_j$,

(3) the number of arguments for which $m_l = 0$ is j,

(4) at least one $m_j > 0$,

(5) the x-sequence has bounded mixed moments of order $4(d-1)$, and

(6) the mixing constants $\alpha_m = O(m^{-2(j+\epsilon)})$ $(\epsilon > 0)$

then the cumulant is $O(n^{p-j})$.

This is an extension of Theorem 2.5. The corollary above can be similarly generalized, but we shall not do so.

2.12 Remark. Given a finite number l of asymptotically stationary sequences, $x^{(1)}, \ldots, x^{(l)}$, we may take $S^{(j)} = \sum_{i=1}^{n} x_i^{(j)}$ and obtain all the analogous results for mixed cumulants provided the vector process $\{(x_i^{(1)}, \ldots, x_i^{(l)}) : i > 0\}$ satisfies the mixing and moment conditions.

We will come across some estimates of the form

$$\sum_{i=1}^{n} \left(\frac{S}{n}\right)^{p_i} \left(\frac{x_n}{n}\right)^{q_i}.$$

Here we are simplifying slightly, because the sums S/n and the factors x_n/n could be from different sequences, and furthermore we might actually have products of factors of the form x_{n-i_j}/n, with a different j for each factor. A primary example is the estimate of covariance. We have

$$
\begin{aligned}
\hat{R}_1 &= \frac{1}{n}\sum_{1}^{n-1}(x_i - \bar{x}_n)(x_{i+1} - \bar{x}_n) \\
&= \frac{1}{n}\sum_{1}^{n-1}(x_i - \mu)(x_{i+1} - \mu) \\
&\quad - \frac{1}{n}(x_n - \mu)(x_{n+1} - \mu) \\
&\quad - (\bar{x}_n - \mu)^2\left(1 + \frac{1}{n}\right) \\
&\quad + \frac{(\bar{x}_n - \mu)}{n}\left[(x_1 - \mu) + (x_n - \mu)\right].
\end{aligned} \tag{1}
$$

There are two sequences involved here: $(x_i - \mu)$ and $(x_i - \mu)(x_{i+1} - \mu)$. Each term does have the form suggested above, except that there may be an extra factor of n^{-p}. We want to conclude that the cumulants of such expressions have the same asymptotic behavior as do the Taylor expansions dealt with in Theorem 2.5.

Theorem 2.5, after rescaling, shows that $\sum_{m=0}^{k} \alpha_m (S/n)^m$ has cumulants κ_j of order $O(n^{1-j})$. This jth cumulant is a linear combination of terms

$$\mathrm{cum}(S^{p_1}, \ldots, S^{p_j}).$$

Subject to appropriate mixing and moment conditions, Remark 2.11 shows that if an $(x_n/n)^q$ argument is adjoined, then the order of the cumulant is actually reduced. Alternatively, if a term S^{p_j} is multiplied by $(x_n/n)^q$ for integer $q > 0$, then the order of the cumulant is unchanged if $q = 1$, and otherwise reduced. Therefore we have the following theorem.

2.13 Theorem. *Let $f_n = \sum_{m=0}^{k}(S/n)^m(x_n/n)^{q_m}\alpha_m$, where the q_m are nonnegative integers and $S = \sum_1^n x_i$. Let $p = j\cdot[\max_m(m+q_m)] > 1$. Assume*

(1) the x's have uniformly bounded mixed moments of order $4(p-1)$, and

(2) the x's are mixing with mixing constants $\alpha_m = O(m^{-2(j+\epsilon)})$.

Then the jth cumulant of f_n is $O(n^{1-j})$ as $n \to \infty$. □

2.14 Remark. Suppose a finite number of sequences $x^{(1)}, \ldots, x^{(r)}$ satisfy the vector versions of (1) and (2) in the last theorem. Here (1) means that the moments of form $Ex_{i_1}^{(m_1)}\cdots x_{i_l}^{(m_l)}$ are uniformly bounded for $l \le 4(p-1)$, and (2) becomes the obvious extension of the mixing concept. We may then replace any factors (S/n) by any $n^{-1}\sum_1^n x_i^{(m)}$, and any x_n/n by $x_{n-i}^{(m)}/n$ for fixed i, or by $x_i^{(m)}/n$ for fixed i. Also, dividing some terms by additional factors of n certainly will not increase the order.

Let us illustrate how the estimate $\hat{R}_1$ of equation (2.12.1) fits this form. Let $x^{(1)}$ be the sequence $(x_i - \mu)$ and $x^{(2)}$ be the sequence $(x_i - \mu)(x_{i+1} - \mu)$. Then the standard estimate of (2.12.1) is

$$\frac{S^{(2)}}{n} - \frac{1}{n}x_n^{(2)} - \left(\frac{S^{(1)}}{n}\right)^2\left(1 + \frac{1}{n}\right) + \frac{S^{(1)}}{n}\left(\frac{x_1^{(1)}}{n} + \frac{x_n^{(1)}}{n}\right).$$

We can therefore conclude that the jth cumulant of $\hat{R}_1$ is

$O(n^{1-j})$. The maximum number of x factors in any term of $\hat{R}_1$ is two, so it will suffice to have bounded mixed x moments of order $4(2j-1)$ and mixing constants $\alpha_m = O(m^{-2(j+\epsilon)})$.

Moment identities

2.15 Conditions. The identities in this section require several types of conditions on the sequence $\{x_i : i \geq 1\}$. The parameters p and l below are specified in each subsection.

(1) Mixing. The x's should have mixing constants $\alpha_n = O(n^{-2p})$.

(2) Moment condition. Mixed x-moments $E|x_{i_1} \cdots x_{i_j}|$ are uniformly bounded for all integer j less than or equal to the order of the moment being considered.

(3) Asymptotic stationarity. If E denotes expectation and E_s denotes expectation with respect to the limiting stationary distribution (assumed to exist), then $|E_s x_{i_1} \cdots x_{i_m} - E x_{i_1} \cdots x_{i_m}| = O(n^{-p})$ whenever the smallest index i_j is at least n and the number of indexes (m) is no more than l.

(4) The asymptotic mean is zero.

2.16 Notation. We will be expressing quantities like $E(\sum_1^n x_i)^p$ as polynomials in n, and we need to establish a notation for the coefficients. "Δ" in the notation below is meant to indicate that the asymptotic means have been subtracted off, so in effect the x's are assumed to have asymptotic mean zero. If S_x denotes $\sum_1^n x_i$, we will write, for example,

$$E\left(\sum_1^n x_i\right) = \mu^0(\Delta S_x)n^0 + O(f(n))$$

$$E\left(\sum_1^n x_i\right)^2 = \mu^1(\Delta S_x{}^2)n^1 + \mu^0(\Delta S_x{}^2)n^0 + O(f(n))$$

$$E\left(\sum_1^n x_i\right)^2\left(\sum_1^n y_i\right) = \mu^1(\Delta S_x{}^2\Delta S_y)n^1 + \mu^0(\Delta S_x{}^2\Delta S_y)n^0 + O(f(n))$$

$$E(\hat{R}_j - R_j)^4 = [\mu^{-2}(\Delta R_j{}^4)n^{-2} + \mu^{-3}(\Delta R_j{}^4)n^{-3} + \mu^{-4}(\Delta R_j{}^4)n^{-4} + O(f(n))]$$

and so on. The $O(f(n))$ above is a generic error term which will be specified in what follows.

Thus the superscript of μ corresponds to the power of n, and the argument of $\mu^p(\cdot)$ is the quantity whose moments are being considered. The notation, then, decomposes the expectation into its components.

We will write μ_s to refer to expectations with respect to the limiting stationary distribution. In addition, if $\hat{R}$ is a vector of estimates and a is a vector, $\mu^p(\Delta R^T a)$ is $\sum \mu^p(\Delta R_i a_i)$, and a similar notation applies to matrix estimates $\hat{M}$.

If the context is clear, we may drop the argument: μ^p. The hatted notation $\hat{\mu}^p(\cdot)$ refers to the estimate of the un-hatted quantity.

We may also need to consider

$$E\left(\sum_{1}^{n} x_i\right)^p \left(\sum_{n+1}^{2n} x_i\right)^q$$

and

$$E\left(\sum_{1}^{n} x_i\right)^p (x_{n+1})^q.$$

Fortunately, in such contexts the underlying sequence x_i will be evident, and we will use the same notation as before but with arguments $(\Delta S_1^p \Delta S_2^q)$ and $(\Delta S_x^p \Delta x_{n+1}^q)$, respectively.

2.17 First moments. Evidently,

(1)
$$\mu_s^0(\Delta S_x) = 0$$
$$\mu^0(\Delta S_x) = \sum_{1}^{\infty} Ex_i,$$

with error term of order $O(n^{-(p-1)})$, provided $p > 1$ in the conditions of Subsection 2.15.

2.18 Second moments. Given that $S_n = \sum_1^n x_i$, and R_i is the correlation of the x sequence at lag i, in the stationary case we get

$$E(S_n)^2 = \sum_{k=0}^{n-1}\left(ES_{k+1}^2 - ES_k^2\right)$$
$$= \sum_{k=0}^{n-1}\left(Ex_{k+1}^2 + 2E\sum_{i=1}^{k} x_i x_{k+1}\right)$$

$$= \sum_{k=0}^{n-1}\left(R_0 + \sum_{1}^{\infty} R_i - 2\sum_{k+1}^{\infty} R_i\right)$$

$$= n\left(R_0 + 2\sum_{1}^{\infty} R_i\right) - 2\sum_{i}^{\infty} iR_i + 2\sum_{i=n+1}^{\infty}(i-n)R_i,$$

provided R_n is $O(n^{-p})$ for $p > 2$, in which case the remainder term (the last term above) is $O(n^{-(p-2)})$.

The asymptotically stationary case is similar, but there is an additional term

$$\sum_{k=1}^{n}\sum_{i=1}^{n} Ex_ix_k - E_sx_ix_k.$$

The subscripted E_s denotes expectation with respect to the asymptotically stationary distribution.

From this,

$$\mu^1(\Delta S_n) = \mu_s^1(\Delta S_n) = R_0 + 2\sum_{1}^{\infty} R_i$$

$$\mu_s^0(\Delta S_n) = -2\sum_{1}^{\infty} iR_i \tag{1}$$

$$\mu^0(\Delta S_n) = \mu_s^0(\Delta S_n) + \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} Ex_ix_k - E_sx_ix_k.$$

To estimate the error term in the above, we need to bound the tail of the double sequence, which is

$$\left|\sum_{1}^{\infty}\sum_{1}^{\infty} Ex_ix_j - E_sx_ix_j - \sum_{1}^{n}\sum_{1}^{n} Ex_ix_j - E_sx_ix_j\right|.$$

This is bounded by

$$2\sum_{i=[n/2]}^{\infty}\sum_{j=1}^{i}|Ex_ix_{i+j}-E_sx_ix_{i+j}|+|Ex_{i+j}x_j-E_sx_{i+j}x_j|,$$

which is $O(n^{-(p-2)})$ for $p>2$ and $l\geq 8$ in the conditions of Subsection 2.15. (We use the notation $[x]$ for the greatest integer less than or equal to x, and we also use the notation $\lceil x\rceil$ for the smallest integer greater than or equal to x).

2.19 Third moments. Here we will indicate the general method and the results, but omit computational details. Write

$$(1)\qquad \begin{aligned} ES_n^3 &= \sum_{k=0}^{n-1} ES_{k+1}^3 - ES_k^3 \\ &= \sum_{k=0}^{n-1} Ex_{k+1}^3 + 3Ex_{k+1}^2\sum_1^k x_i + 3Ex_{k+1}\left(\sum_1^k x_i\right)^2. \end{aligned}$$

We now rewrite non-stationary items in terms of stationary ones, plus an error term. Thus

$$\sum_0^{n-1} Ex_{k+1}^3 = \sum_0^{n-1}\left[E_sx_{k+1}^3 + \left(Ex_{k+1}^3 - E_sx_{k+1}^3\right)\right].$$

Other terms in (1) may be reduced by splitting sums:

$$\sum_1^k = \sum_1^{[k/2]} + \sum_{[k/2]+1}^{k}.$$

Applying these techniques leads to

$$\mu^1(\Delta S_n{}^3) = \sum_{i,j=-\infty}^{\infty} E_s x_0 x_i x_j + 3\mu^0(\Delta S_n)\mu^1_s(\Delta S_n{}^2).$$

The conditions are that $p > 2$, and $l \geq 8$ in the conditions of Subsection 2.15. We will not need $\mu^0(\Delta S_n{}^3)$.

2.20 Fourth moments. We now can begin to use the cumulant bounds developed in the first part of this chapter. If x_n satisfies the hypotheses of Theorem 2.5 for fourth order cumulants, we can expand the fourth cumulant of $S = \sum_1^n x_i$ in terms of moments:

$$\begin{aligned} O(n) &= \mathrm{cum}(S, S, S, S) \\ &= ES^4 - 4ES^3ES - 3E^2S^2 + 12ES^2E^2S - 6E^4S. \end{aligned}$$

Equating the $O(n^2)$ terms of the above shows that

$$\mu^2(\Delta S_n{}^4) = 3\left(\mu^1(\Delta S_n{}^2)\right)^2.$$

In view of the more general results of the first part of this chapter, we can certainly draw a similar conclusion for quantities like $\hat{R}_j$. We will also need to know $\mu^1(\Delta S_n{}^4)$. In the context of autoregressive processes, we will have an easy way of estimating the coefficient in the stationary case, but we will need the correction for the asymptotically stationary case. The method of derivation is the same as that indicated in the last subsection. After some algebra we find

$$\mu^1(\Delta S_n{}^4) = -3\left(\mu^1_s(\Delta S_n{}^2)\right)^2 + E_s x^4 + 4\mu^0_s(\Delta S\,\Delta x^3_{n+1})$$

$$+ 6\mu_s^0(\Delta S^2 \Delta x_{n+1}^2) + 4\mu_s^0(\Delta S^3 \Delta x_{n+1}^1)$$
$$+ 4\mu^0(\Delta S_n)\mu_s^1({\Delta S_n}^3)$$
$$+ 6\left[\mu^0({\Delta S_n}^2) - \mu_s^0({\Delta S_n}^2)\right]\mu_s^1({\Delta S_n}^2).$$

The last two lines above gives the correction for the asymptotically stationary case. For the coefficients of this subsection to be valid, it suffices that $l \geq 12$ and $p > 4$ in Subsection 2.15.

2.21 Fifth and sixth order moments. If x_n satisfies the hypotheses of Theorem 2.5 for fifth order cumulants, then

$$\mu^2({\Delta S_n}^5) = 10\mu^1({\Delta S_n}^3)\mu^1({\Delta S_n}^2) - 15\left(\mu^1(\Delta S_n)\right)^2\mu^0(\Delta S_n)$$
$$\mu^3({\Delta S_n}^6) = 15\left(\mu^1({\Delta S_n}^2)\right)^3.$$

It is fortunate that not only are the most significant terms of the higher order moments reduced to coefficients of lower order moments, but also that the nonstationary correction is based only on first and second moment nonstationary terms.

In the next subsection, we extend these formulas to the case of mixed moments. These results will be very useful in subsequent calculations.

2.22 Mixed moments. Given two or more sequences, $\{a_i : i \geq 1\}$, $\{b_i : i \geq 1\}$, and so on, we may easily extend the moment identities of the preceding paragraphs by identifying the coefficients of

$$E\left(s\sum a_i + t\sum b_i\right)^p,$$

provided $sa_i + tb_i$ satisfies the sufficient conditions of Subsection 2.15. This gives us the formulas enumerated below, where $S_a = \sum_1^n a_i$, and similarly for S_b and S_c.

In many cases, a "μ" may be replaced by a "μ_s." Those formulas derived using Theorem 2.5 (which means the fourth order and higher ones) will apply to non-sample means like $\hat{R}_j$.

$$\mu^1(\Delta S_a \Delta S_b) = E_s a_0 b_0 + \sum_1^\infty (E_s a_0 b_i + E_s a_i b_0) \tag{1}$$

$$\begin{aligned}\mu^0(\Delta S_a \Delta S_b) = &- \sum_1^\infty k(E_s a_0 b_k + E_s a_k b_0) \\ &+ \sum_{i,j=1}^\infty (E a_i b_j - E_s a_i b_j)\end{aligned} \tag{2}$$

$$\begin{aligned}\mu^1(\Delta S_a{}^2 \Delta S_b) = &\sum_{i,j=-\infty}^\infty E_s a_0 a_i b_j \\ &+ 2\mu^0(\Delta S_a)\mu^1(\Delta S_a \Delta S_b) \\ &+ \mu_s^1(\Delta S_a{}^2)\mu^0(\Delta S_b)\end{aligned} \tag{3}$$

$$\begin{aligned}\mu^2(\Delta S_a{}^2 \Delta S_b \Delta S_c) = &\mu_s^1(\Delta S_a{}^2)\mu_s^1(\Delta S_b \Delta S_c) \\ &+ 2\mu_s^1(\Delta S_a \Delta S_b)\mu_s^1(\Delta S_a \Delta S_c)\end{aligned} \tag{4}$$

$$\begin{aligned}\mu^2(\Delta S_a{}^4 \Delta S_b) = &4\mu^1(\Delta S_a{}^3)\mu^1(\Delta S_a \Delta S_b) \\ &+ 6\mu^1(\Delta S_a{}^2 \Delta S_b)\mu^1(\Delta S_a{}^2) \\ &- 3\left(\mu^1(\Delta S_a{}^2)\right)^2 \mu^0(\Delta S_b) \\ &- 12\mu^1(\Delta S_a{}^2)\mu^1(\Delta S_a \Delta S_b)\mu^0(\Delta S_a)\end{aligned} \tag{5}$$

$$(6)\qquad \mu^3\left(\Delta S_a{}^4 \Delta S_b \Delta S_c\right) = 3\left(\mu^1\left(\Delta S_a{}^2\right)\right)^2 \mu^1(\Delta S_b \Delta S_c) + 12\mu^1\left(\Delta S_a{}^2\right)\mu^1(\Delta S_a \Delta S_b)\mu^1(\Delta S_a \Delta S_c).$$

3

Summing difference recursions

In our moment calculations we will have to compute sums like

$$\sum_{0}^{\infty} y_i, \quad \sum_{0}^{\infty} i y_i, \quad \text{or} \quad \sum_{i=0}^{\infty} w_{i,i+j}$$

where the y's satisfy a difference equation and the w's satisfy a difference equation in each index.

Sequences indexed by non-negative integers

3.1 One dimensional recursions. We assume that the sequence $\{y_i : i \geq 0\}$ satisfies a strictly stable order k difference equation,

$$y_n + a_1 y_{n-1} + \cdots a_k y_{n-k} = 0$$

for $n \geq k$, where $y_0, \ldots, y_{k-1}$ are known. That the recursion is strictly stable means that all the roots of the characteristic polynomial $z^k + a_1 z^{k-1} + \cdots + a_k z^0$ lie in the interior of the unit circle. Under this assumption, the sums in question converge absolutely.

(See PRIESTLEY (1981) for further discussion and results). We may calculate these sums directly, or by using the generating function $\varphi(z) = \sum_0^\infty y_j z^j$. Then $\varphi(1) = \sum_0^\infty y_i$ and $\varphi'(1) = \sum_1^\infty j y_j$.

Because the y_i's satisfy the homogeneous difference equation, we find

$$\varphi(z)(1 + \cdots + a_k z^k) = y_0 + z(y_1 + a_1 y_0) + \cdots + z^{k-1}(y_{k-1} + \cdots + a_{k-1} y_0)$$

and

$$\sum_0^\infty y_i = \varphi(1)$$

$$(1) \qquad = \frac{y_0(1 + \cdots + a_{k-1}) + \cdots + y_{k-1}(1)}{(1 + \cdots + a_k)}.$$

It is possible to calculate $\sum j y_j$ by considering $\varphi'(1)$, but practically speaking there is an easier method. We have

$$\sum_1^\infty j y_j = \sum_{j=1}^\infty y_j \sum_{i=1}^j 1$$

$$= \sum_{i=1}^\infty \sum_{j=i}^\infty y_j.$$

Note that the sums $S_i = \sum_i^\infty y_j$ obey the same difference recursion as the y_i's, but with different initial values. Therefore a convenient method to compute this sum is as follows.

3.2 Algorithm. To compute $\sum_1^\infty j y_j$:

(1) Compute $S_0 = \sum_0^\infty y_i$ from (1) above.

(2) Compute $S_1 = S_0 - y_0, \ldots, S_k = S_{k-1} - y_k$.

(3) The answer is then

$$\frac{S_1(1 + \cdots + a_{k-1}) + \cdots + S_k(1)}{(1 + \cdots + a_k)}.$$

$\square$

3.3 Two dimensional recursions. We will also need to sum two dimensional arrays whose elements satisfy a difference equation. Specifically, if R_{ij} is the non-stationary expectation $E(x_i - \mu)(x_j - \mu)$ and R_i is the stationary expectation $E_s(x_i - \mu)(x_0 - \mu)$, we will need to sum the array $R_{ij} - R_{|i-j|}$ in its entirety and also along its diagonals.

In an autoregressive model of order k in which the sequence $\{x_i - \mu : i \geq 1\}$ satisfies the usual difference equation and $y_i = x_i - \mu$, we get

$$\begin{aligned} R_{ij} &= Ey_iy_j \\ &= -Ey_i(a_1y_{j-1} + \cdots + a_ky_{j-k} + \epsilon_j) \\ &= -\left(a_1R_{i,j-1} + \cdots + a_kR_{i,j-k} + Ey_i\epsilon_j\right), \end{aligned}$$

assuming $j > k$. This gives R_{ij} as the solution of a non-homogeneous difference equation. But y_i is a linear combination of the first k y_i's and of $\epsilon_{k+1}, \ldots, \epsilon_i$, and therefore $Ey_i\epsilon_j$ is some multiple of $E\epsilon_j^2$. In particular, $Ey_i\epsilon_j$ will be the same whether we regard $y_1, \ldots, y_k$ as fixed or as coming from the stationary distribution. This means that $R_{|i-j|}$ is a particular solution of the non-homogeneous difference equation, and so $R_{ij} - R_{|i-j|}$ satisfies the homogenous difference equation

$$R_{ij} - R_{|i-j|} = \delta_{i,j}$$
$$\triangleq -\left(a_1\delta_{i,j-1} + \cdots + a_k\delta_{i,j-k}\right).$$

The problems of summing the entire array and of suming a diagonal require different approaches.

Summing the entire array is very simple. Just use formula (3.1.1) to sum rows zero through $k-1$ of the array. Next, input these sums as initial values in (3.1.1) once more to obtain the sum of the entire array. In our applications, the double sequences of interest will satisfy the same difference equation in both indexes, but the above procedure works in general.

Next assume we want to compute $\sum_{i=0}^{\infty} w_{i,i+j}$, where the double sequence w satisfies the same order k difference equation in both indexes $(i, j \geq 0)$. The difficulty is that the one dimensional sequence $w_{i,i+j}$ (for fixed j) does not satisfy an order k difference equation. We will see, however, that an order k vector difference equation is satisfied.

Let $v_\alpha = [y_{\alpha,\alpha}, \ldots, y_{\alpha,\alpha+k-1}]^T$ (the superscript T indicates transpose), and let F be a $k \times k$ matrix defined by the requirement that $F[y_{\alpha-k}, \ldots, y_{\alpha-1}]^T = [y_{\alpha-k+1}, \ldots, y_\alpha]$. Note that to calculate Fv for any vector v we simply shift the components up, calculate the bottom component from the difference equation, and discard the original top component. In other words, we need not store the entire matrix F in the computer, and furthermore multiplication by F is much easier than usual matrix multiplication.

Now, from the difference equation and the definition of F we know

$$(1) \qquad v_\alpha + Fv_{\alpha-1}a_1 + \cdots + F^k v_{\alpha-k}a_k = 0.$$

This is the required vector difference equation. From this we deduce

$$\sum_{\alpha=0}^{\infty} v_\alpha = (I + a_1 F + \cdots + a_k F^k)^{-1}$$

$$\times \left[v_0 + (v_1 + Fv_0 a_1) + \cdots + (v_{k-1} + \cdots F^{k-1} v_0 a_{k-1}) \right].$$

Let $A(z) = a_k z^k + \cdots + a_0 z^0$. It is easy to show that the eigenvalues of F are $r_1, \ldots, r_k$ with eigenvectors $[1, \ldots, r_i^{k-1}]^T$, where the r_i are the roots of $A(z^{-1}) = 0$, which are all less than one in absolute value. (If the roots are not distinct we can draw the same conclusions using a limiting argument). Therefore

$$I + a_1 F + \cdots + a_k F^k = \prod_1^k (I - r_i F),$$

where each factor is invertible since the matrix norm $\|r_i F\|_2$ is less than one. In other words, the inverse in the expression for $\sum_0^\infty v_\alpha$ exists.

In applications of the above formula, we will not generally know $v_1, \ldots, v_{k-1}$. Let b_i be defined as $[y_{i,0}, \ldots, y_{i,k-1}]^T$, for $i = 0, \ldots, k-$

1. These b's will be known and because $v_\alpha = F^\alpha b_\alpha$, we may rewrite (1) as

$$\sum_{\alpha=0}^{\infty} v_\alpha = (I + a_1 F + \cdots + a_k F^k)^{-1}$$

$$\times \left[b_0 + F(b_1 + b_0 a_1) + \cdots + F^{k-1}(b_{k-1} + \cdots + b_0 a_{k-1}) \right].$$

The matrix polynomials may be efficiently computed in nested form ("Horner's rule"), which has the added advantage that we only multiply by F, which is easy.

For convenience we will refer to the function which takes as input $y_0, \ldots, y_{k-1}$ and produces $\sum_0^\infty y_i$ as output as

$$\mathrm{SUM}_{01}(y_0, \ldots, y_{k-1}),$$

and that which produces $\sum_1^\infty y_i$ as $\mathrm{SUM}_{11}(y_0, \ldots, y_{k-1})$. The first subscript (0 or 1) gives the starting index for the summation, and the second indicates the univariate case. Similarly,

$$\mathrm{SUM}_{02}(b_0, \ldots, b_{k-1}) = \sum_0^\infty v_\alpha$$

and

$$\mathrm{SUM}_{12}(b_0, \ldots, b_{k-1}) = \sum_1^\infty v_\alpha.$$

Difference recursions indexed by $(\ldots, -1, 0, 1, \ldots)$

3.4 An example. We want to extend our arsenal of summation methods. The technique we are about to develop will greatly simplify certain calculations. It is best illustrated with a simple example for which it is not really necessary.

Suppose that a sequence $\{y_i : i \geq 0\}$ satisfies the usual stable autoregressive difference equation of order k and is moreover stationary (that is, $y_0, \ldots, y_{k-1}$ are chosen according to the limiting distribution). We want to calculate $\sum_{i=-\infty}^{\infty} \text{cum}(y_0, y_i)$, which is the same as $\sum_{-\infty}^{\infty} R_i$, where R_i is the covariance of the process at lag i. The covariances R_i do obey a difference equation, and the straightforward method to compute the sum is to partition it into two sums, $\sum_0^{\infty} + \sum_{-\infty}^{-1}$. The latter sum is the same as $\sum_1^{\infty}$.

Though easy to do in this case, splitting the sum would be much more difficult if we were dealing with

$$\sum_{i,j=-\infty}^{\infty} \text{cum}(y_0 y_l, y_i y_j).$$

For one thing, we would need to estimate k^3 fourth order cumulants, where k is the order of the autoregression. This is analagous to the need to know $\hat{R}_0, \ldots, \hat{R}_{k-1}$ in order to compute $\sum_0^{\infty} \hat{R}_i$. Furthermore, although the sum $\sum_{i,j=-\infty}^{\infty}$ involves only two indexes, some of the component sums will involve three indexes after shifting. For example, the sum over $(i \geq 0, j \leq 0)$, after adding j to each index to shift the indexes to non-negative values involves three indexes. This means we would in effect have to compute three dimensional sums of the form

$$\sum_{i=0}^{\infty} w_{i,l+i,j+i}$$

where w obeys a difference equation in each index. This all can be done, but it is not the best way.

Let us write a formal z-transform:

$$\varphi(z_1, z_2) = \sum_{j,l=-\infty}^{\infty} \operatorname{cum}(y_j, y_l) z_1^j z_2^l. \tag{1}$$

We say "formal" because this sum does not converge for any (z_1, z_2). Ignoring all such convergence problems for the moment, we could recover $\operatorname{cum}(y_j, y_l)$ from the double contour integral (z_1 and z_2 each take values on the unit circle):

$$\frac{1}{(2\pi i)^2} \int \frac{\varphi(z_1, z_2)}{z_1^{j+1} z_2^{l+1}} d(z_1, z_2).$$

If $A(z) = a_k z^k + \cdots + a_0 z^0$, we can "solve" for φ by multiplying both sides of (1) by $A(z_1)A(z_2)$, where the a's are the coefficients of the difference equation:

$$a_0 y_j + \cdots + a_k y_{j-k} = \epsilon_j,$$

which we assume to hold for all integer j, positive and negative. Equation (1) becomes

$$A(z_1)A(z_2)\varphi(z_1, z_2) = \sum_{i,j=-\infty}^{\infty} \operatorname{cum}(\epsilon_j, \epsilon_l) z_1^j z_2^l.$$

Because of the assumed independence of the ϵ's, this is

$$\sum_{m=-\infty}^{\infty} \sigma_\epsilon^2 z_1^m z_2^m.$$

Substituting back into the inversion integral for $\mathrm{cum}(y_j, y_l)$ converts the double contour integral into the sum of the products of two countour integrals, namely

$$\sigma_\epsilon^2 \sum_{m=-\infty}^{\infty} \left(\frac{1}{2\pi i} \int \frac{z_1^m}{A(z_1) z_1^{j+1}} dz_1 \right) \left(\frac{1}{2\pi i} \int \frac{z_2^m}{A(z_2) z_2^{l+1}} dz_2 \right).$$

These contour integrals are equal to w_{j-m} and w_{l-m} respectively, where w_i is 0 for negative indexes and satisfies the difference equation with coefficients $a_0, \ldots, a_k$ subject to initial conditions $w_0 = 1, (w_1 + a_1 w_0) = \cdots = (w_{k-1} + \cdots + a_{k-1} w_0) = 0$ (as follows by writing the z-transform of this w). This suggests

$$\mathrm{cum}(y_j, y_l) = \sum_{m=-\infty}^{\infty} w_{j-m} w_{l-m}.$$

This conclusion is in fact valid, as we will show later. Then

$$\sum_{l=-\infty}^{\infty} \mathrm{cum}(y_0, y_l) = \sum_{l,m=-\infty}^{\infty} w_{-m} w_{l-m} \sigma_\epsilon^2.$$

Note that the sum with respect to m is a two dimensional recursion, but the sum with respect to l is one dimensional, and is the same for all m. Summing first over l and then over m gives

$$\sum_{l=-\infty}^{\infty} \mathrm{cum}(y_0, y_l) = \sigma_\epsilon^2 \left(\sum_0^{\infty} w_i \right)^2.$$

Even the sum $\sum_0^\infty w_i$ is easy to compute. It is $1/A(1)$, and so

$$\sum_{l=-\infty}^{\infty} \operatorname{cum}(y_0, y_l) = \frac{\sigma_\epsilon^2}{A(1)^2}.$$

When this general method is applied to higher order quantities, the only new parameters to be estimated will be third and fourth moments of ϵ, as opposed to $\operatorname{cum}(y_0, y_i, y_j, y_l)$ for all i, j, and l between 0 and $k-1$. Furthermore, we will only have to sum at most two dimensional recursions.

3.5 Validity of the method. Let $\{Y_{i,j} : (i,j) \in \mathbf{Z}^2\}$ be a double sequence on $\mathbf{Z}^2$ which satisfies a difference equation in each index. The difference equation is not assumed to be homogeneous. In the example of the previous subsection, $\operatorname{cum}(y_i, y_j)$ was such a $Y_{i,j}$. Let B_1 be the linear operator which shifts the first index, and let B_2 be that which shifts the second: B_1 shifts $Y_{i,j}$ into position $(i+1, j)$. If $\delta^{i,j}$ is the double sequence which is identically zero except that entry (i,j) is one, then $B_1 \delta^{i,j} = \delta^{i+1,j}$.

Multiplying the formal z-transform of the double sequence of cumulants $\{\operatorname{cum}(y_j, y_l)\}$ by the product $A(z_1)A(z_2)$ to get the formal z-transform of cumulants $\{\epsilon_{i,j}\} \triangleq \{\operatorname{cum}(\epsilon_i, \epsilon_j)\}$ corresponds to the following linear operation:

$$\Lambda Y \triangleq A(B_1)A(B_2)Y = \epsilon.$$

The linear operator $A(B_1)$ is by definition of the polynomial A equal to $I + a_1 B_1 + \cdots + a_k B_1^k$. In the example of the last subsection,

$\epsilon = \sum_{m=-\infty}^{\infty} \sigma_\epsilon^2 \delta^{m,m}$. We need to show, in some sense, that the bounded linear functional Λ is invertible with a continuous inverse. Because $A(B_1)A(B_2)$ can be factored into terms $(I - r_i B_1)$ and $(I - r_i B_2)$ with $|r_i| < 1$, it will suffice to assume $\Lambda = I - rB$, where $|r| < 1$ and B is either B_1 or B_2.

Define the measure $\nu\{l\} = (1 + l^2)^{-1}$ for $l \in \mathbf{Z}$, and $\mu(l,m) = \nu\{l\}\nu\{m\}$. Consider $L^p(\mu)$ on $\mathbf{Z}^2$, for $p \in (1,\infty)$. The map Λ is a bounded linear functional on $L^p(\mu)$, and furthermore it is one to one. To see this, observe that $\Lambda Y = 0$ and $Y \neq 0$ together imply that $Y_{j,l}$ grows exponentially fast as $j \to -\infty$, so $Y \notin L^p(\mu)$. (This might not be the case had we chosen μ to decay exponentially instead of polynomially).

Next we want to find the inverse $\Lambda^{-1} = (I - rB)^{-1}$. The obvious candidate is

$$\Lambda^{-1} = I + rB + r^2 B^2 + \cdots.$$

We can bound the operator norm $\|B^k\|$ $(= \sup \|Bx\|/\|x\|)$ by $3k^2$, so in fact the series defining Λ^{-1} is Cauchy and the candidate inverse is well defined and a bounded linear functional (hence continuous). Furthermore,

$$\begin{aligned} \Lambda^{-1}\Lambda Y &= \lim_{l\to\infty} (I + rB + \cdots + r^l B^l)(I - rB)Y \\ &= \lim_{l\to\infty} (Y - r^{l+1} B^{l+1} Y) \\ &= Y. \end{aligned}$$

In our applications, Λ will be the composite of two or more maps of the form $A(B)$, and ΛY will have a simple form. For example, in the last subsection $\Lambda Y = \epsilon = \sum_{-\infty}^{\infty} \sigma_\epsilon^2 \delta^{m,m}$. The inverse of $\delta^{0,0}$ is $w = \{w_i w_j\}$, where w_i is zero for negative indexes, $w_0 = 1$, and $(w_1 + a_1 w_0) = \cdots = (w_{k-1} + \cdots + a_{k-1} w_0) = 0$. By linearity, $\Lambda^{-1}\delta^{\alpha,\beta}$ has (j,l) element $w_{j-\alpha} w_{l-\beta}$. In our example, $\sum_{-\infty}^{\infty} \delta^{m,m}\sigma_\epsilon^2$ converges to ϵ in $L^p(\mu)$, so continuity of Λ^{-1} guarantees that $Y_{l,j} = \sum_{m=-\infty}^{\infty} \sigma_\epsilon^2 w_{l-m} w_{j-m}$. (Being the $L^p(\mu)$ limit in this case implies being the pointwise limit). Because the same holds for any order of summation of $\sum_{-\infty}^{\infty} \delta^{m,m}\sigma_\epsilon^2$, the pointwise convergence is absolute. In general, when Y is bounded $\Lambda Y = \epsilon$ will be bounded, so $\sum \delta^{\alpha,\beta}\epsilon_{\alpha,\beta}$ will converge to ϵ in $L^p(\mu)$. We have proved the following, which clearly extends to any number of indexes.

3.6 Theorem. *Let $z^k + \cdots + a_k z^0$ have all its roots in the interior of the unit circle, and let $A(B)$ denote the operator $(I + \cdots + a_k B^k)$. Let Y be a bounded sequence on $\mathbf{Z}^2$, and $\epsilon = A(B_1)A(B_2)Y$, where B_1 and B_2 are shift operators on the first and second indexes as in the previous discussion. Then*

$$Y_{l,j} = \sum_{\alpha,\beta=-\infty}^{\infty} \epsilon_{\alpha,\beta} w_{l-\alpha} w_{j-\beta},$$

where w is as in the previous paragraph. The convergence is absolute. □

4

Derivation of corrections

Introduction

We will be using Cornish-Fisher expansions as discussed in the *Introduction to derive more accurate* confidence intervals for the autoregressive process $x = \{x_i : i \geq 1\}$ which satisfies our usual stable difference equation

$$(x_i - \mu) + \cdots + a_k(x_{j-k} - \mu) = \epsilon_j.$$

We assume the difference equation is strictly stable, which means that the roots of the characteristic polynomial $z^k + \cdots + a_k z^0$ all lie in the interior of the unit circle. In our derivation, we take as given the validity of the Cornish-Fisher expansion and of the methods developed in the preceding chapters. Refer to the articles by TANIGUCHI (1984), ABRAMOVITCH AND SINGH (1985), and GOTZE AND HIPP (1983) for a further discussion of sufficient conditions for the validity of Cornish-Fisher and Edgeworth Expansions.

We will assume that the errors ϵ_i are independent and identically distributed, have a Lebesgue density component which is positive on an interval, and have finite moments of all orders. (Recall that a general distribution may be expressed as the sum of a distribution absolutely continuous with respect to Lebesgue measure, a discrete distribution, and a singular distribution. The first of these is the Lebesgue density component).

First we will form a zero order pivot, which is the usual pivot t_n given by

$$t_n = \frac{\sqrt{n}(\bar{x} - \mu)}{\sqrt{v_n}}.$$

Here, we are given $x_1, \ldots, x_n$; $\bar{x}_n$ is the sample mean, which is an estimate of the asymptotic mean $\lim_{i\to\infty} Ex_i$ denoted by μ; and v_n is the estimate of the variance constant v for $\bar{x}_n$, $v \triangleq \lim_{n\to\infty} nE(\bar{x}_n - \mu)^2$. Our first order correction, T_1, is a Cornish-Fisher expansion based on t_n:

$$T_1 = t_n + \hat{\theta} n^{-1/2} + \hat{\rho} t_n^2 n^{-1/2}.$$

The Cornish-Fisher expansion as given in KENDALL AND STUART (1977) is

$$T_1 = t_n - \kappa_1 - \frac{1}{6}\kappa_3(t_n^2 - 1),$$

where κ_i refers to the ith cumulant of t_n. Our correction is the same, except that we must estimate the cumulants. That we estimate instead of using the true values does not matter to the level of accuracy required for a first order correction—see ABRAMOVITCH AND SINGH (1985).

The next step is to form a Cornish-Fisher expansion from T_1, where now the κ_i's refer to cumulants of T_1:

$$T_2 = T_1 - \kappa_1 - \frac{1}{6}(T_1^2 - 1) - \frac{1}{2}(\kappa_2 - 1)T_1 + \frac{1}{3}\kappa_1\kappa_3 T_1 - \frac{1}{24}\kappa_4(T_1^3 - 3T_1) + \frac{1}{36}\kappa_3^2(4T_1^3 - 7T_1).$$

(See KENDALL AND STUART (1977) or HILL AND DAVIS (1968) for the statement and derivation of the Cornish Fisher expansion). T_2 differs from a standard normal random variable by $o_p(n^{-1})$. It will be possible to obtain t_n as a cubic polynomial in T_2, (which we take to be standard normal) instead of vice versa. This will make the formation of confidence intervals easier.

The zero order pivot

4.1 Algorithm. We begin by specifying a possible pivot t_n, which will be used in what follows. Let $t_n = n^{1/2}v_n^{-1/2}(\bar{x}_n - \mu)$, where the estimate v_n of the variance constant for $\bar{x}_n$ is obtained as follows.

1. Estimate the asymptotic mean μ by $\hat{\mu} = \bar{x}_n = n^{-1}\sum_1^n x_i$.
2. Estimate covariances R_j for $j = 0, \dots, k$ by

$$\begin{aligned}\hat{R}_j &= \frac{1}{n}\sum_{i=1}^{n-j}(x_i - \bar{x}_n)(x_{i+j} - \bar{x}_n)\\ &= \frac{1}{n}\left[\sum_1^{n-j}(x_i - \mu + \mu - \bar{x}_n)(x_{i+j} - \mu + \mu - \bar{x}_n)\right]\end{aligned}$$

$$= \frac{1}{n}\sum_{1}^{n-j}(x_i - \mu)(x_{i+j} - \mu) + \frac{n-j}{n}(\bar{x}_n - \mu)^2$$

$$+ \frac{(\mu - \bar{x}_n)}{n}\left[\sum_{i=1}^{n-j}(x_{i+j} - \mu) + \sum_{i=1}^{n-j}(x_i - \mu)\right]$$

$$= \frac{1}{n}\sum_{1}^{n-j}(x_i - \mu)(x_{i+j} - \mu) + \frac{n-j}{n}(\bar{x}_n - \mu)^2$$

$$- \frac{(\bar{x}_n - \mu)}{n}\left[2n(\bar{x}_n - \mu) - \right.$$

$$\left. \sum_{1}^{j}(x_i - \mu) - \sum_{n-j+1}^{n}(x_i - \mu)\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)(x_{i+j} - \mu) \tag{1}$$

$$- \frac{1}{n}\sum_{i=n-j+1}^{n}(x_i - \mu)(x_{i+j} - \mu)$$

$$- (\bar{x}_n - \mu)^2\left(1 + \frac{j}{n}\right)$$

$$+ \frac{(\bar{x}_n - \mu)}{n}\left[\sum_{1}^{j}(x_i - \mu) + \sum_{n-j+1}^{n}(x_i - \mu)\right].$$

The last form is convenient for analysis, while in practice $\hat{R}_j$ would be calculated by a method like the fast Fourier transform.

3. Solve the Yule-Walker equations for the estimated autore-

gressive coefficients $\hat{a}^T = [\hat{a}_1, \ldots, \hat{a}_k]^T$:

$$\begin{bmatrix} \hat{R}_0 & \cdots & \hat{R}_{k-1} \\ \vdots & \ddots & \vdots \\ \hat{R}_{k-1} & \cdots & \hat{R}_0 \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_k \end{bmatrix} = - \begin{bmatrix} \hat{R}_1 \\ \vdots \\ \hat{R}_k \end{bmatrix}.$$

Or, denoting the above matrices by $\hat{M}$, $\hat{a}$, and $-\hat{R}$,

$$\hat{M}\hat{a} = -\hat{R}.$$

4. Compute the estimate $\hat{\sigma}^2$ of the prediction error variance (variance of the ϵ's of the model):

$$\hat{\sigma}_n^2 \triangleq \hat{\sigma}^2 = 1 \cdot \hat{R}_0 + \cdots + \hat{a}_k \hat{R}_k$$
$$= \hat{R}_0 + \hat{a}^T \hat{R}.$$

5. Estimate the variance constant v by

$$v_n = \frac{\hat{\sigma}_n^2}{(1 + \hat{a}_1 + \cdots + \hat{a}_k)^2}.$$

6. The pivot is $t_n = n^{1/2} v_n^{-1/2} (\bar{x}_n - \mu)$.

The first order pivot

4.2 Introduction. In order to make a first order correction to the confidence interval for the stationary mean $\mu = E_s x$, we need to estimate Et_n and Et_n^3, where t_n is the zero order pivot. Specifically, if $Et_n = \alpha_1 n^{-1/2} + \alpha_2 n^{-1} + O(n^{-3/2})$ and $Et_n^3 = \beta_1 n^{-1/2} + \beta_2 n^{-1} + O(n^{-3/2})$, the first order pivot T_1 is given by

$$T_1 = t_n + \hat{\theta} n^{-1/2} + \hat{\rho} t_n^2 n^{-1/2},$$

where $\theta = -3\alpha_1/2 + \beta_1/6$ and $\rho = \alpha_1/2 - \beta_1/6$. This follows from the form of the Cornish-Fisher expansion given at the beginning of the chapter along with the relationships between moments (μ_i) and cumulants (κ_i):

$$\kappa_1 = \mu_1,$$

$$\kappa_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3.$$

In deriving T_1, we will ignore terms which are $O_p(n^{-1})$. (The definition of "$O_p(n^{-1})$" is similar to that of "$o_p(n^{-1})$": a sequence $\{z_i : i \geq 1\}$ is said to be $O_p(n^{-1})$ if nz_n is *bounded* in probability).

4.3 First moment of t_n. From a Taylor expansion,

$$t_n = n^{-1/2}(\bar{x}_n - \mu)\left(v^{-1/2} - \frac{1}{2}v^{-3/2}(v_n - v)\right) + O_p(n^{-1}).$$

Then given our regularity conditions stated in the first paragraph of this chapter ($x = \{x_i : i \geq 1\}$ satisfies a stable order k autoregressive difference equation whose errors ϵ_i are independent and identically distributed with positive Lebesgue density component on an interval and with moments of all orders),

$$Et_n = En^{-1/2}(\bar{x}_n - \mu)v^{-1/2}$$
$$- \frac{1}{2}En^{-1/2}v^{-3/2}(v_n - v)(\bar{x}_n - \mu) + O(n^{-1}).$$

The regularity conditions imply that $E(\bar{x}_n - \mu) = n^{-1}\mu^{-1}(\Delta\bar{x}) + O(n^{-3/2})$, where as usual $\mu^{-1}(\Delta\bar{x})$ means $\sum_1^\infty E(x_i - \mu)$. Therefore the first element of the $n^{-1/2}$ coefficient of Et_n (that is, α_1) is simply

$\mu^{-1}(\Delta\bar{x}_n)v^{-1/2}$. Our analysis is conditional on $x_1, \ldots, x_k$, (k is the order of the difference equation) and thus from the formulas for summing difference recursions,

$$\begin{aligned} v^{-1/2}\mu^{-1}(\Delta\bar{x}) &= v^{-1/2}\mathrm{SUM}_{01}(x_1 - \mu, \ldots, x_k - \mu) \\ &= v^{-1/2}\frac{(x_1 - \mu)(1 + \cdots + a_{k-1}) + \cdots + (x_k - \mu)\cdot 1}{(1 + \cdots + a_k)}, \end{aligned}$$

and the natural estimate is

$$v_n^{-1/2}\frac{(x_1 - \bar{x}_n)(1 + \cdots + \hat{a}_{k-1}) + \cdots + (x_k - \bar{x}_n)\cdot 1}{(1 + \cdots + \hat{a}_k)}.$$

The other element of α_1 is

$$v^{-3/2}\mu_s^1(\Delta\bar{x}\Delta v) \triangleq En(\bar{x}_n - \mu)(-1/2)v^{-3/2}(v_n - v).$$

Expanding $v_n - v$ as a Taylor series in $(\hat{\sigma}^2 - \sigma^2)$ and $\sum_i(\hat{a}_i - a_i)$ gives the first order approximation

$$\begin{aligned} (1) \qquad v_n - v = &\frac{\hat{\sigma}^2 - \sigma^2}{(1 + \cdots + a_k)^2} \\ &- \frac{2\sigma^2}{(1 + \cdots + a_k)^3}(\hat{a}_1 - a_1 + \cdots + \hat{a}_k - a_k). \end{aligned}$$

The autoregressive coefficients $\hat{a}$ satisfy $\hat{M}\hat{a} = -\hat{R}$, while for the true values we have $Ma = -R$. Writing

$$(M + \Delta M)(a + \Delta a) = -(R + \Delta R)$$

shows that $\Delta Ma + M\Delta a + \Delta M\Delta a = -\Delta R$ or

$$\begin{aligned} \hat{a} - a &\triangleq \Delta a \\ &= -(M + \Delta M)^{-1}(\Delta R + \Delta Ma) \\ &\triangleq -\hat{M}^{-1}(\Delta R + \Delta Ma). \end{aligned}$$

We may similarly find $\Delta(M^{-1}) \triangleq \hat{M}^{-1} - M^{-1}$ by writing

$$\begin{aligned}(M+\Delta M)(M^{-1}+\Delta(M^{-1})) &= I \\ \Rightarrow \Delta(M^{-1}) &= -\hat{M}^{-1}\Delta M M^{-1} \\ &= -M^{-1}\Delta M M^{-1}\end{aligned}$$

to first order. If desired, then, we have a second order approximation for Δa:

$$\begin{aligned}\Delta a &= -M^{-1}(\Delta R + \Delta M a) \\ &\quad + M^{-1}\Delta M M^{-1}(\Delta R + \Delta M a).\end{aligned} \tag{2}$$

For the present, we only need the first term on the right hand side.

The remaining part of the $n^{-1/2}$ coefficient of Et_n (that is, α_1), is then

$$\begin{aligned}&-\frac{1}{2}En(\bar{x}_n-\mu)v^{-3/2}(v_n-v) \\ &\qquad = -\frac{1}{2}v^{-3/2}n\left[\frac{E(\bar{x}_n-\mu)(\hat{\sigma}^2-\sigma^2)}{(1+\cdots+a_k)^2}\right. \\ &\qquad\qquad \left. - \frac{2\sigma^2E(\bar{x}_n-\mu)1^T(\hat{a}-a)}{(1+\cdots+a_k)^3}\right] \\ &\qquad = -\frac{1}{2}v^{-3/2}n\left[\frac{E(\bar{x}_n-\mu)(\hat{\sigma}^2-\sigma^2)}{(1+\cdots+a_k)^2}\right. \\ &\qquad\qquad \left. - \frac{2\sigma^2E(\bar{x}_n-\mu)1^T(-1)M^{-1}(\Delta R+\Delta M a)}{(1+\cdots+a_k)^3}\right].\end{aligned}$$

Recalling that the notation $\mu^{-1}(\Delta\bar{x}\Delta R)$ denotes the most significant vector coefficient of $E(\bar{x}_n-\mu)\Delta R$, with similar definitions for

$\mu^{-1}(\Delta\bar{x}\Delta R_0)$ and the matrix $\mu^{-1}(\Delta\bar{x}\Delta M)$, linearity shows that the formula above is

$$-\frac{1}{2}v^{-3/2}\left[\frac{E(\bar{x}_n-\mu)(\hat{\sigma}^2-\sigma^2)n}{(1+\cdots+a_k)^2}\right.$$
$$\left.+\frac{2\sigma^2}{(1+\cdots+a_k)^3}1^T M^{-1}\left(\mu^{-1}(\Delta x\Delta R)+\mu^{-1}(\Delta x\Delta M)a\right)\right].$$

Next we must express $\hat{\sigma}^2-\sigma^2$. We have

$$\begin{aligned}\hat{\sigma}^2 &= \hat{R}_0+\hat{a}^T R\\ &= (R_0+\Delta R_0)+(a+\Delta a)^T(R+\Delta R),\end{aligned}$$

so to first order

$$\begin{aligned}\Delta\sigma^2 &\triangleq \hat{\sigma}^2-\sigma^2\\ &= \Delta R_0+\Delta a^T R+a^T\Delta R\\ &= \Delta R_0-\left[M^{-1}(\Delta R+\Delta M a)\right]^T R+a^T\Delta R.\end{aligned}$$

Because $a=-M^{-1}R=-M^{-T}R$, this is

$$\begin{aligned}\Delta\sigma^2 &= \Delta R_0+(\Delta R^T+a^T\Delta M)a+a^T\Delta R\\ &= \Delta R_0+2a^T\Delta R+a^T\Delta M a.\end{aligned}$$

Therefore to first order

$$\begin{aligned}\mu^{-1}\big(\Delta\bar{x}\Delta(\sigma^2)\big) = \mu^{-1}(\Delta\bar{x}\Delta R_0)&\\ +2a^T\mu^{-1}(\Delta\bar{x}\Delta R)&\\ +a^T\mu^{-1}(\Delta\bar{x}\Delta M)a.&\end{aligned}$$

Combining results shows

$$-\frac{1}{2}\sqrt{n}v^{-3/2}\hat{\mu}^{-1}(\Delta x\Delta v) =$$

$$-\frac{1}{2}v^{-3/2}\sqrt{n}\left[\frac{\mu^{-1}(\Delta x\Delta R_0)+2a^T\mu^{-1}(\Delta x\Delta R)+a^T\mu^{-1}(\Delta x\Delta M)a}{A(1)^2}\right.$$
$$\left.+\frac{2\sigma^2 1^T M^{-1}(\mu^{-1}(\Delta x\Delta R)+\mu^{-1}(\Delta x\Delta M)a)}{A(1)^3}\right].$$

Therefore, the $n^{-1/2}$ coefficient of Et_n is

$$\text{(3)}\quad \alpha_1 = v_n^{-1/2}\frac{(x_1-\bar{x}_n)(1+\cdots+\hat{a}_{k-1})+\cdots+(x_k-\bar{x}_n)\cdot 1}{(1+\cdots+\hat{a}_k)}$$
$$-\frac{1}{2}v^{-3/2}\sqrt{n}\left[\frac{\mu^{-1}(\Delta x\Delta R_0)+2a^T\mu^{-1}(\Delta x\Delta R)+a^T\mu^{-1}(\Delta x\Delta M)a}{A(1)^2}\right.$$
$$\left.+\frac{2\sigma^2 1^T M^{-1}(\mu^{-1}(\Delta x\Delta R)+\mu^{-1}(\Delta x\Delta M)a)}{A(1)^3}\right].$$

We are left to compute $\mu^{-1}(\Delta x\Delta R_j)$ for $j=0,\ldots,k$. Recall the form of $\hat{R}_j$ given in point 2, equation (4.1.1), in the section on the zero order pivot.

Let $R'_j = n^{-1}\sum_1^n(x_i-\mu)(x_{i+j}-\mu)$. Then

$$\mu^{-1}\left(\Delta x\Delta R'_j\right) = \mu_s^{-1}\left(\Delta x\Delta R'_j\right)$$
$$= \sum_{i=-\infty}^{\infty} E_s(x_0-\mu)(x_j-\mu)(x_i-\mu).$$

The results of Chapter 2 show that $\mu^{-1}(\Delta x\Delta R_j) = \mu^{-1}\left(\Delta x\Delta R'_j\right)$. It is worth noting in passing, however, that the end effects in the computation of $\hat{R}'_j$ can be significant, though they are not in this instance. In particular, the results of Chapter 2 also show that $\mu^{-2}(\Delta x^2\Delta R_j) \neq \mu^{-2}\left(\Delta x^2\Delta R'_j\right)$.

To estimate $\sum_{-\infty}^{\infty} E_s(x_0-\mu)(x_j-\mu)(x_i-\mu)$, we use the methods of Chapter 3. This illustrates the utility of Theorem 3.6. We have

$$E_s(x_l-\mu)(x_j-\mu)(x_i-\mu) = \mathrm{cum}(x_l-\mu, x_j-\mu, x_i-\mu)$$
$$= \sum_{m=-\infty}^{\infty} \mathrm{cum}(\epsilon,\epsilon,\epsilon) y_{m-l} y_{m-j} y_{m-i}$$

where the cumulants are with respect to the stationary distribution and y has z-transform $1/A(z) = (a_0 z^0 + \cdots + a_k z^k)^{-1}$. This means that y_i is zero for negative indexes and obeys the autoregressive difference equation with $y_0 = 1, 0 = y_1 + a_1 y_0 = \cdots = y_{k-1} + \cdots + a_{k-1} y_0$. Therefore

$$\sum_{i=-\infty}^{\infty} E_s(x_0-\mu)(x_j-\mu)(x_i-\mu) =$$
$$\kappa_3(\epsilon) \sum_{i=0}^{\infty} y_i \sum_{m=-\infty}^{\infty} y_{0-m} y_{j-m}$$
$$= \frac{\kappa_3(\epsilon)}{A(1)} \sum_{m=0}^{\infty} y_m y_{m+j}.$$

We already know how to compute $[\sum y_m y_{m+j}]_{j=0}^{k-1}$, because this is just $\mathrm{SUM}_{02}[y_\alpha y_\beta]$. The special structure of $[y_\alpha y_\beta]$ facilitates the calculation of the sum, as reflected in the algorithmic summary at the end of the chapter.

4.4 Third moment of t_n. For the first order pivot T_1 we also need to estimate β_1, the $n^{-1/2}$ coefficient of Et_n^3, or $\mu^{-1/2}(\Delta t_n{}^3)$.

From a Taylor expansion,

$$t_n^3 = n^{3/2}(\bar{x}_n - \mu)^3 \left(v^{-3/2} - \frac{3}{2} v^{-5/2}(\bar{v}_n - v) \right) + O_p(n^{-1}).$$

Therefore $\beta_1 = v^{-3/2}\mu^{-2}(\Delta x^3) - 3v^{-5/2}\mu^{-2}(\Delta x^3 \Delta v)/2$. Now

$$\begin{aligned}
\mu_s^{-2}(\Delta x^3) &= \sum_{i,j=-\infty}^{\infty} E_s(x_0 - \mu)(x_i - \mu)(x_j - \mu) \\
&= \sum_{i,j=-\infty}^{\infty} \kappa_3(\epsilon) \sum_{m=-\infty}^{\infty} y_{-m} y_{i-m} y_{j-m} \\
&= \frac{\kappa_3(\epsilon)}{A(1)^3}.
\end{aligned}$$

From the fact that $\mu^{-2}(\Delta x^3) = \mu_s^{-2}(\Delta x^3) + 3\mu^{-1}(\Delta x)v$, we find

$$\begin{aligned}
\beta_1 &= \frac{\kappa_3(\epsilon) v^{-3/2}}{A(1)^3} + 3\mu^{-1}(\Delta x) v^{-1/2} \\
&\quad - \frac{3}{2} v^{-5/2} \mu^{-2}(\Delta x^3 \Delta v) \\
&= \frac{\kappa_3(\epsilon) v^{-3/2}}{A(1)^3} + 3\mu^{-1}(\Delta x) v^{-1/2} \\
&\quad - \frac{9}{2} v^{-3/2} \mu^{-1}(\Delta x \Delta v).
\end{aligned}$$

Replacing all quantities in the above equation and in equation (4.3.3) for α_1 by their natural estimates enables us to compute

$$\hat{\theta} = -\frac{3}{2}\hat{\alpha}_1 + \frac{1}{6}\hat{\beta}_1,$$

$$\hat{\rho} = \frac{1}{2}\hat{\alpha}_1 - \frac{1}{6}\hat{\beta}_1$$

$$T_1 = t_n + \frac{\hat{\theta}}{\sqrt{n}} + \frac{\hat{\rho} t_n^3}{\sqrt{n}}.$$

The second order pivot

The second order pivot T_2 derives from a Cornish Fisher expansion of T_1. The expansion is

$$\begin{aligned} T_2 = T_1 - \kappa_1 - \frac{1}{6}\kappa_3(T_1^2 - 1) - \frac{1}{2}(\kappa_2 - 1)T_1 \\ + \frac{1}{3}\kappa_1\kappa_3 T_1 - \frac{1}{24}\kappa_4(T_1^3 - 3T_1) \\ + \frac{1}{36}\kappa_3^2(4T_1^3 - 7T_1) + o_p(n^{-1}), \end{aligned}$$

where κ_i is the ith cumulant of T_1. It is not hard to show that κ_1 and κ_3 are $o(n^{-1})$, while κ_2 and κ_4 are $O(n^{-1})$. Because $t_n - T_1$ is $O_p(n^{-1/2})$, we will rewrite the above as

$$\begin{aligned} T_2 = T_1 - \frac{1}{2}(\kappa_2 - 1)t_n \\ - \frac{1}{24}\kappa_4(t_n^3 - 3t_n) + o_p(n^{-1}). \end{aligned}$$

Again, the fact that we estimate $n\kappa_2$ and $n\kappa_4$ does not matter to $o_p(n^{-1})$—see ABRAMOVITCH AND SINGH (1985).

4.5 Second cumulant of T_1. We have

$$\begin{aligned} \kappa_2 - 1 &= ET_1^2 - E^2T_1 - 1 \\ &= ET_1^2 - 1 + o(n^{-1}) \\ &= E\left(t_n + \frac{\theta_n}{\sqrt{n}} + \frac{\rho_n t_n^2}{\sqrt{n}}\right)^2 - 1 + o(n^{-1}) \\ &= Et_n^2 - 1 + \frac{\theta^2}{n} + \frac{3\rho^2}{n} + \frac{2\theta\alpha_1}{n} \end{aligned}$$

$$+\frac{2\rho\beta_1}{n}+\frac{2\theta\rho}{n}$$

$$+\frac{2E(\rho_n-\rho)t_n^3}{\sqrt{n}}+\frac{2Et_n(\theta_n-\theta)}{\sqrt{n}}+o(n^{-1}).$$

Estimation of most of these quantities, and of those needed for κ_4, is straightforward. The necessary calculations are summarized in the subsection at the end of the chapter, along with a few brief explanations and references to Chapters 2 and 3. In the current subsection, we will discuss points which require further elaboration, and will point out some pitfalls.

In many cases we will need the most significant coefficient of covariance between the point estimate $\bar{x}_n$ and the product of two estimates, $\hat{p}\hat{q}$. In other words, we want the most significant part of $E(\bar{x}_n-\mu)(\hat{p}\hat{q}-pq)$. To "first order" the change $\hat{p}\hat{q}-pq$ is $p\Delta q+q\Delta p$, so we expect the coefficient to be

$$p\mu^{-1}(\Delta x\Delta q)+q\mu^{-1}(\Delta x\Delta p).$$

This conclusion is generally correct, but one must be a little careful of the logic. The true change includes a $\Delta p\Delta q$ term, but the expectation $E(\Delta x\Delta p\Delta q)$ is of lower order than the other terms. If, however, we were considering $E(\bar{x}_n-\mu)^2(\hat{p}\hat{q}-pq)$, we would need to account for the extra $E\Delta x^2\Delta p\Delta q$ term, because this is of the same order as $E\Delta x^2\Delta p$ and $E\Delta x^2\Delta q$. In short, one must carefully apply the results of Chapter 2.

We now turn to specific items, following the same order as in the second order summary at the end of the chapter.

1. $\hat{\mu}_s^{-2}\left(\Delta\bar{x}_n^2\Delta R'_j\right)$. From the definition of R'_j and the mixed third moment identities, this is

$$\sum_{l,m=-\infty}^{\infty} \operatorname{cum}((x_0-\mu)(x_j-\mu),(x_l-\mu)(x_m-\mu)).$$

Again we will see the usefulness of Theorem 3.6.

Let $y_{i,j,l,m} = \operatorname{cum}((x_i-\mu)(x_j-\mu),(x_l-\mu)(x_m-\mu))$. Applying the autoregressive difference equation to each coordinate transforms this sequence to the following ϵ sequence:

$$\begin{aligned}\epsilon_{i,j,l,m} &= \operatorname{cum}(\epsilon_i\epsilon_j,\epsilon_l\epsilon_m)\\ &= \begin{cases} E\epsilon^4-\sigma^4 & \text{for } i=j=l=m,\\ \sigma^4 & \text{for } i=l\neq j=m,\\ \sigma^4 & \text{for } i=m\neq j=l,\\ 0 & \text{otherwise.}\end{cases}\end{aligned}$$

Theorem 3.6 then shows

$$\begin{aligned} y_{i,j,l,m} = \left(E\epsilon^4-3\sigma^4\right)&\sum_{p=-\infty}^{\infty} y_{i-p}y_{j-p}y_{l-p}y_{m-p}\\ &+\sigma^4\sum_{p,q=-\infty}^{\infty} y_{i-p}y_{j-q}y_{l-p}y_{m-q}\\ &+\sigma^4\sum_{p,q=-\infty}^{\infty} y_{i-p}y_{j-q}y_{l-q}y_{m-p}.\end{aligned}$$

Recall that y is the sequence whose z-transform is $1/A(1)$. Including those pairs (p,q) for which $p=q$ in the latter two sums is compensated by the $-3\sigma^4$ in the first. Summing the above over $(l,m)\in\mathbf{Z}^2$ now leads to the result stated in the summary.

2. $\hat{\mu}_s^{-2}\left(\Delta\bar{x}_n^2\Delta R_j\right)$. The expression of the summary is obvious, except for the $-j\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n^2\Delta R_j\right)$ term. This comes from

$$-E(\bar{x}_n-\mu)^2\cdot\frac{1}{n}\sum_{n-j+1}^{n}(x_i-\mu)(x_{i+j}-\mu).$$

To find the value of this, it is only necessary to equate $O(n^2)$ terms of

$$O(n)=\mathrm{cum}\left(\textstyle\sum_1^n(x_i-\mu),\sum_1^n(x_i-\mu),(x_{n-i}-\mu),(x_{n-i+j}-\mu)\right).$$

3. $\hat{\mu}_s^{-1}\left(\Delta R_j'\Delta R_l'\right)$. Again use Theorem 3.6. From the moment identities, we want to estimate

$$\sum_{m=-\infty}^{\infty}\mathrm{cum}((x_0-\mu)(x_j-\mu),(x_m-\mu)(x_{l+m}-\mu))$$

under the stationary distribution. Theorem 3.6 reduces part of this to a sum

$$\sum_{m,p=-\infty}^{\infty}y_{-p}y_{j-p}y_{m-p}y_{l+m-p}.$$

This accounts for the S_jS_l term appearing in the summary. The rest is

$$\sum_{m,p,q=-\infty}^{\infty}y_{-p}y_{j-q}y_{m-p}y_{l+m-q}$$
$$+\sum_{m,p,q=-\infty}^{\infty}y_{-p}y_{j-q}y_{m-q}y_{l+m-p}.$$

We will illustrate the method for the first of these two sums. First sum over p and then over q. The result is

$$\sum_{m=-\infty}^{\infty} S_{|m|}S_{|l+m-j|}.$$

For convenience suppose $l \geq j$. Then both $\sum_0^\infty$ and $\sum_{-\infty}^{-|l-j|}$ are equal to

$$S_{2,|l-j|} \triangleq \sum_{m=0}^{\infty} S_m S_{m+|l-j|}.$$

In general, this leads to

$$2S_{2,|l-j|} + \sum_{0<m<|l-j|} S_m S_{|l-j|-m},$$

except that if $l = j$ we need to subtract S_0^2 from this (because then $\sum_0^\infty$ and $\sum_{-\infty}^{-|l-j|}$ overlap).

4. $\hat{\mu}^{-2}(\Delta\bar{x}_n^2\Delta a)$. Note that we consider the entire expression for Δa given in the first order correction, Formula (4.3.2). In general,we will need to consider higher order errors to compute $\hat{\mu}^{-2}(\Delta\bar{x}_n^2\Delta(\cdot))$ or $\hat{\mu}^{-1}(\Delta(\cdot))$ for any "$(\cdot)$."

5. $\hat{\mu}^{-1}(\Delta x\Delta(\epsilon^3))$. The estimate of $E\epsilon^3$ is given by

$$\begin{aligned}\hat{E}\epsilon^3 &= n^{-1}\sum_{k+1}^{n} \hat{\epsilon}_i^3 \\ &= n^{-1}\sum_{k+1}^{n} \left[(x_i - \bar{x}_n) + \cdots + \hat{a}_k(x_{i-k} - \bar{x}_n)\right]^3.\end{aligned}$$

This reduces to our problem to estimating the most significant coefficient of

$$E(\bar{x}_n - \mu)n^{-1}\sum_1^n \epsilon_i^3$$

and

$$E(\bar{x}_n - \mu)n^{-1}\sum_1^n (\hat{\epsilon}_i^3 - \epsilon_i^3).$$

The first of these coefficients we refer to as $\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n\Delta(n^{-1}\sum\epsilon_i^3\right)$. Because $\hat{\epsilon}^3 - \epsilon^3$ is a polynomial, we could write the second exactly as a Taylor series, but we only need the following approximation:

$$E(\bar{x}_n - \mu)n^{-1}\sum_1^n (\hat{\epsilon}_i^3 - \epsilon_i^3) =$$

$$E(\bar{x}_n - \mu)\left[3n^{-1}\sum_1^n \epsilon_i^2\left[-A(1)(\bar{x}_n - \mu) + \sum_{j=1}^k \Delta a_j(x_{i-j} - \mu)\right]\right].$$

Part of this is, to our order of aproximation,

$$-3A(1)\sigma^2 E(\bar{x}_n - \mu)^2.$$

The rest is

$$\sum_{j=1}^k E\Delta a_j(\bar{x}_n - \mu)3n^{-1}\sum_{i=1}^n \epsilon_i^2(x_{i-j} - \mu).$$

The inner sum estimates $\mathrm{cum}\left(\epsilon_i^2, x_{i-j}\right)$, which is zero. The expectation $E\Delta a_j(\bar{x}_n - \mu)\Delta\left(\mathrm{cum}\left(\epsilon_i^2, x_{i-j}\right)\right)$ is of lower order, though of course the estimate need not be zero.

6. Calculation of M_2 and $\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta M_2)$. The algorithm of the summary uses Horner's rule or nested evaluation of the matrix polynomial defining M_2. From the stated algorithm, it's clear that M_2 is computed as

$$F\left(\cdots\left(F\hat{a}_k I + \hat{a}_{k-1} I\right) + \cdots\right) + a_0 I.$$

The calculation of $\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta M_2)$ parallels this.

The rest of the calculation of $\hat{\kappa}_2(T_1)$ is straightforward.

4.6 Fourth cumulant of T_1.

1. $\hat{\mu}_s^{-3}(\Delta\bar{x}_n^4)$. From Theorem 3.6, the fourth order cumulant of the sum $\sum_1^n (x_i - \mu)$ under the stationary distribution is $\kappa_4(\epsilon)/A(1)^4$. Now equate the $O(n)$ terms of

$$\kappa_4 = \mu_4 - 3\mu_2{}^2,$$

where the κ and μ's are moments and cumulants of $\sum_1^n (x_i - \mu)$ under the stationary distribution.

2. $\hat{\mu}^{-1}(\Delta R_j)$. By this we mean the n^{-1} coefficient of $E(\hat{R}_j - R_j)$. Note that in general in computing $\hat{\mu}^{-2}(\Delta(\cdot))$ we must account for higher order terms. For example, to compute $\hat{\mu}^{-1}(\Delta a)$ we use the first and second order terms of Δa given in the first order correction.

3. $\hat{\kappa}_4(T_1)$. The expression given derives from the relationship between non-central moments and cumulants,

$$\kappa_4 = \mu_4 - 4\mu_3\mu_1 - 3\mu_2{}^2 + 12\mu_2\mu_1{}^2 - 6\mu_1{}^4,$$

which in the case of T_1 reduces to

$$ET_1^4 - 3 - 6\left(\kappa_2(T_1) - 1\right).$$

Inverting the correction

We now discuss how to invert the first or second order correction. Suppose

$$\xi = t_n + \frac{\theta_n}{\sqrt{n}} + \frac{\rho_n t_n^2}{\sqrt{n}} + \frac{\nu_n t_n}{n} + \frac{\omega_n t_n^3}{n},$$

and define

$$g(x, \theta, \rho, \nu, \omega) = -\left(\frac{\theta}{\sqrt{n}} + \frac{\rho x^2}{\sqrt{n}} + \frac{\nu x}{n} + \frac{\omega x^3}{n}\right).$$

By θ_n (and so on) we mean the estimate of θ based on $x_1, \ldots, x_n$. Then $g(t_n, \theta_n, \rho_n, \nu_n, \omega_n)$ is $O_p(n^{-1})$, as are $g(\xi, \theta_n, \rho_n, \nu_n, \omega_n)$ and $t_n - \xi$. Ignoring terms which are $O_p(n^{-3/2})$ and writing g for the function $g(\xi, \theta_n, \rho_n, \nu_n, \omega_n)$ and g' for the derivative of g with respect to its first argument,

$$\begin{aligned}
t_n - \xi &= g + (t_n - \xi)g' + \frac{1}{2}(t_n - \xi)^2 g'' + O_p(n^{-3/2}) \\
&= g + g' \cdot (g + (t_n - \xi)g') + O_p(n^{-3/2}) \\
&= g + g'g + O_p(n^{-3/2}) \\
&= -\left(\frac{\theta_n}{\sqrt{n}} + \frac{\rho_n \xi^2}{\sqrt{n}} + \frac{\nu_n \xi}{n} + \frac{\omega_n \xi^3}{n}\right) \\
&\qquad + \frac{1}{n}\left(2\theta_n \rho_n \xi + 2\rho_n^2 \xi^3\right) + O_p(n^{-3/2}).
\end{aligned}$$

Therefore

$$t_n = \xi - \frac{1}{\sqrt{n}}\left(\theta_n + \rho_n \xi^2\right) + \frac{1}{n}\left(-\nu_n \xi + 2\theta_n \rho_n \xi - \omega_n \xi^3 + 2\rho_n^2 \xi^3\right) + O_p(n^{-3/2}).$$

Here ξ is regarded as a normal random variable. This equation enables us to convert normal quantiles (ξ values) into t_n quantiles directly, rather than solving for the t_n quantiles.

Algorithmic summary

In this section, we summarize the computations necessary to make the first and second order corrections. As usual, x_i is an autoregressive process of order k. The autoregressive coefficients are $a_0 \triangleq 1, a_1, \ldots, a_k$, the errors are $\epsilon_k, \ldots, \epsilon_n$, the variance of the ϵ's is σ^2, while v is the asymptotic variance constant for $\bar{x}_n$. The analysis is conditional on $x_1, \ldots, x_k$. The first and second order corrections depend on the method of estimating v, and we assume that the method detailed at the beginning of the chapter is used.

4.7 First order summary.

1. Compute

$$\mu^{-1}(\Delta \bar{x}_n) = \mathrm{SUM}_{01}(x_1 - \bar{x}_n, \ldots, x_k - \bar{x}_n).$$

2. Estimate ϵ_i, $i = k+1, \ldots, n$. The estimates are

$$\hat{\epsilon}_i = (x_i - \bar{x}_n) + \hat{a}_1(x_{i-1} - \bar{x}_n) \cdots + \hat{a}_k(x_{i-k} - \bar{x}_n).$$

3. Calculate $\hat{E}\epsilon^3 = n^{-1} \sum_{k+1}^{n} \hat{\epsilon}_i^3$.

4. Calculate $1/\hat{A}(1) = (1 + \cdots + \hat{a}_k)^{-1}$. This is the sum $\sum_0^\infty y_i$, where y_i has the z-transform $1/\hat{A}(z)$. Recall that $A(z) = a_k z^k + \cdots + a_0 z^0 \quad (a_0 = 1)$.

5. Calculate $\sum_{i=0}^{\infty} y_i y_{i+j}$, $(j = 0, \ldots, k-1)$ as follows. First solve the triangular system

$$\begin{bmatrix} \hat{a}_0 & & 0 \\ \vdots & \ddots & \\ \hat{a}_{k-1} & \ldots & \hat{a}_0 \end{bmatrix} \begin{bmatrix} y_0 \\ \vdots \\ y_{k-1} \end{bmatrix} = e_1 \triangleq \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}.$$

Recall that F is a shift matrix defined by the requirement that

$$F \begin{bmatrix} y_0 \\ \vdots \\ y_{k-1} \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix}.$$

The matrix F, then, involves the estimated autoregressive coefficients. The desired sums are the solution of

$$(I + \cdots + \hat{a}_k F^k) \begin{bmatrix} \sum y_\alpha^2 \\ \vdots \\ \sum y_\alpha y_{\alpha+k-1} \end{bmatrix} = \begin{bmatrix} y_0 \\ \vdots \\ y_{k-1} \end{bmatrix}.$$

6. Compute $\sum y_\alpha y_{\alpha+k}$ from the sums of the previous paragraph using the fact that the sums $\sum_{\alpha \geq 0} y_\alpha y_{\alpha+i}$ as a function of i obey the same order k difference equation.

7. Compute the covariances between the point estimate $\bar{x}_n$ and estimated covariances $\hat{R}_j$:

$$\hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta R_j) = \frac{\hat{E}\epsilon^3}{\hat{A}(1)} \sum_{\alpha \geq 0} y_\alpha y_{\alpha+j},$$

for $j = 0, \ldots, k$.

8. Let $\hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta M)$ denote the symmetric $k \times k$ Toeplitz matrix with (i,j) entry equal to $\hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta R_{|i-j|}\big)$. Similarly let $\hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta R)$ be the column vector with ith entry equal to $\hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta R_i)$, for i between 1 and k.

9. Compute the estimated coefficient of covariance between the point estimate $\bar{x}_n$ and the vector of estimated autoregressive coefficients $\hat{a}$:

$$\hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta a) = -\hat{M}^{-1}\left(\hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta R) + \hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta M)\hat{a}\right),$$

and also

$$\hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta(1^T a)\big) = 1^T \hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta a).$$

10. Calculate the estimated coefficient of covariance between the point estimate and prediction error variance estimate:

$$\hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta(\sigma^2)\big) = \hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta R_0) + \hat{R}^T \hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta a) + a^T \hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta R).$$

11. Calculate the estimated coefficient of covariance between the point estimate and v_n :

$$\hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta v) = \frac{\hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta(\sigma^2)\big)}{(1 + \cdots + \hat{a}_k)^2} - \frac{2\hat{\sigma}^2 \hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta(1^T a)\big)}{(1 + \cdots + \hat{a}_k)^3}.$$

12. The estimate of α_1, that is, the $n^{-1/2}$ coefficient of Et_n, is given by

$$\hat{\alpha}_1 = \hat{\mu}^{-1}(\Delta \bar{x}_n) v_n^{-1/2} - \frac{1}{2}\hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta v) v_n^{-3/2}.$$

13. The most significant coefficient of the third central moment of the point estimate, under the stationary distribution is then:

$$\hat{\mu}_s^{-2}(\Delta\bar{x}_n^3) = \frac{\hat{E}\epsilon^3}{\hat{A}(1)^3}.$$

14. The estimated $n^{-1/2}$ coefficient of Et_n^3 is then

$$\hat{\beta}_1 = \hat{\mu}_s^{-2}(\Delta\bar{x}_n^3)v_n^{-3/2} + 3\hat{\mu}^{-1}(\Delta\bar{x}_n)v_n^{-1/2}$$
$$- \frac{9}{2}\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta v)v_n^{-3/2}.$$

15. One may now compute the coefficients for the Cornish-Fisher polynomial:

$$\hat{\theta} = -\frac{3}{2}\hat{\alpha}_1 + \frac{1}{6}\hat{\beta}_1,$$
$$\hat{\rho} = \frac{1}{2}\hat{\alpha}_1 - \frac{1}{6}\hat{\beta}_1.$$

16. The first order corrected pivot is then

$$T_1 = t_n + \hat{\theta}n^{-1/2} + \hat{\rho}t_n^2 n^{-1/2},$$

and one may compute the p-quantile of the distribution of t_n by

$$z_p - \hat{\theta}n^{-1/2} - \hat{\rho}z_p^2 n^{-1/2},$$

where z_p is the p-quantile of the standard normal distribution.

4.8 Second order summary. The bulk of this calculation consists of the calculation of the second cumulant of T_1. Though it may appear lengthy, the calculation is not too computer-intensive.

Calculate $\hat{\mu}^{-2}(\Delta x^2)$ using the moment identities:

1. $$\sum_1^\infty \hat{R}_i = \mathrm{SUM}_{11}\left(\hat{R}_0, \ldots, \hat{R}_{k-1}\right),$$

2. $$\hat{\mu}_s^{-2}(\Delta \bar{x}_n^2) = -2 \cdot \mathrm{SUM}_{01}\left(\textstyle\sum_1^\infty \hat{R}_i, \ldots, \sum_k^\infty \hat{R}_i\right).$$

Estimate the difference between the above and its non-stationary equivalent. From Chapter 2, this amounts to summing the double array δ below over the entire first quadrant.

3. $$\delta_{i,j} = (x_i - \bar{x}_n)(x_j - \bar{x}_n) - \hat{R}_{|i-j|},$$
$$\text{for } 1 \leq i, j \leq k.$$

4. $$\mathrm{TEMP}_i = \mathrm{SUM}_{01}(\delta_{i,1}, \ldots, \delta_{i,k}),$$
$$\text{for } i = 1, \ldots, k.$$

5. $$\hat{\mu}^{-2}(\Delta \bar{x}_n^2) - \hat{\mu}_s^{-2}(\Delta \bar{x}_n^2) =$$
$$\mathrm{SUM}_{01}(\mathrm{TEMP}_1, \ldots, \mathrm{TEMP}_k).$$

6. $$\hat{\mu}^{-2}(\Delta \bar{x}_n^2) = \hat{\mu}_s^{-2}(\Delta \bar{x}_n^2) + \left(\hat{\mu}^{-2}(\Delta \bar{x}_n^2) - \hat{\mu}_s^{-2}(\Delta \bar{x}_n^2)\right).$$

In the next sequence of equations we will compute $\hat{\mu}^{-2}(\Delta \bar{x}_n^2 \Delta R_j)$. Recall that R'_j is $n^{-1}\sum(x_i - \mu)(x_{i+j} - \mu)$.

7. $$\hat{E}\epsilon^4 = \frac{1}{n}\sum_{k+1}^n \hat{\epsilon}_i^4,$$

8. $$\hat{\mu}_s^{-2}\left(\Delta \bar{x}_n^2 \Delta R'_j\right) = \frac{\hat{E}\epsilon^4 - 3\hat{\sigma}^4}{\hat{A}(1)^2} \sum_{\alpha \geq 0} y_\alpha y_{\alpha+j}$$
$$+ \frac{2\hat{\sigma}^4}{\hat{A}(1)^4}.$$

Refer to the first order correction for $\sum y_\alpha y_{\alpha+j}$, which has already been computed.

9. $$\hat{\mu}^{-1}\left(\Delta R_j'\right) = \mathrm{SUM}_{02}\left[(x_i - \bar{x}_n)(x_l - \bar{x}_n) - R_{|i-l|}\right]_j,$$

where the value for $j = k$ is obtained from the values for $j = 0, \ldots, k-1$, using the fact that as a function of j, these quantities obey the autoregressive difference equation. Next, for $j = 0, \ldots, k$, compute

10. $$\begin{aligned}\hat{\mu}_s^{-2}\left(\Delta\bar{x}_n^2 \Delta R_j\right) &= -3\left(\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n^2\right)\right)^2, \\ &\quad - j\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n^2\right)R_j + \hat{\mu}_s^{-2}\left(\Delta\bar{x}_n^2 \Delta R_j'\right),\end{aligned}$$

11. $$\begin{aligned}\hat{\mu}^{-2}\left(\Delta\bar{x}_n^2 \Delta R_j\right) &= \hat{\mu}_s^{-2}\left(\Delta\bar{x}_n^2 \Delta R_j\right) \\ &\quad + 2\hat{\mu}^{-1}\left(\Delta\bar{x}_n\right)\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n \Delta R_j'\right) \\ &\quad + \hat{\mu}_s^{-1}\left(\Delta\bar{x}_n^2\right)\hat{\mu}^{-1}\left(\Delta R_j'\right).\end{aligned}$$

In the following formulas, S_i will stand for $\sum_{\alpha\geq 0} y_\alpha y_{\alpha+i}$, which has been computed in the first order correction. Let $S_{2,|l-j|}$ stand for the sum $\sum_{\alpha\geq 0} S_\alpha S_{\alpha+|l-j|}$. Compute:

12. $$S_{2,|l-j|} = \mathrm{SUM}_{02}(S_\alpha S_\beta)_{|l-j|},$$

and iterate the above out to $S_{2,2k}$ using the autoregressive difference equation (which $S_{2,i}$ obeys as a function of i). Next estimate $\mu_s^{-1}\left(\Delta R_j' \Delta R_l'\right)$:

13. $$\mathrm{TEMP} = \left(\hat{E}\epsilon^4 - 3\hat{\sigma}^4\right) S_j S_l$$

$$+\hat{\sigma}^4\Bigg[2S_{2,|l-j|}+2S_{2,l+j}$$

$$+\sum_{0<i<|l-j|} S_i S_{|l-j|-i}$$

$$+\sum_{0<i<l+j} S_i S_{l+j-i}\Bigg].$$

The estimate of $\mu_s^{-1}\left(\Delta R'_j \Delta R'_l\right)$ is then

14. $$\hat{\mu}_s^{-1}\left(\Delta R'_j \Delta R'_l\right)=\begin{cases}\text{TEMP}-S_0^2\hat{\sigma}^4, & \text{if } l=j\neq 0\\ \text{TEMP}-2S_0^2\hat{\sigma}^4, & \text{if } l=j=0\\ \text{TEMP}, & \text{otherwise.}\end{cases}$$

15. $$\hat{\mu}_s^{-1}(\Delta R_j \Delta R_l)=\hat{\mu}_s^{-1}\left(\Delta R'_j \Delta R'_l\right).$$

In order to estimate $\mu^{-2}\left(\Delta \bar{x}_n^2 \Delta a\right)$, first compute the following:

16. $$\hat{\mu}_s^{-2}(\Delta \bar{x}_n^2 \Delta R_j \Delta R_l)=$$
$$\hat{\mu}_s^{-1}(\Delta \bar{x}_n^2)\hat{\mu}_s^{-1}(\Delta R_j \Delta R_l)$$
$$+2\hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta R_j)\hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta R_l).$$

Let $M^{(j)}$ be a $k\times k$ matrix whose (i,l) entry is one if $|i-l|=j$ and zero otherwise; for $i\leq j\leq k$ let $R^{(j)}$ be a column vector with elements $1,\ldots,k$ which are zero except for the jth element which is one; and let $R^{(0)}=0$. Compute:

17. $$W_j^{(1)}=\hat{M}^{-1}(R^{(j)}+M^{(j)}\hat{a}),$$

18. $$W_{i,j}^{(2)}=\hat{M}^{-1}M^{(i)}W_j^{(1)},$$

19. $$\hat{\mu}_s^{-1}(\Delta a \Delta R_l)=-\sum_{j=0}^{k} W_j^{(1)}\hat{\mu}_s^{-1}(\Delta R_l \Delta R_j),$$

for $l = 0, \ldots, k$.

20. $$\hat{\mu}^{-2}(\Delta\bar{x}_n^2 \Delta a) = \sum_{0 \le i,j \le k} \hat{\mu}_s^{-2}(\Delta\bar{x}_n^2 \Delta R_i \Delta R_j) W_{i,j}^{(2)} - \hat{M}^{-1}\left(\hat{\mu}^{-2}(\Delta\bar{x}_n^2 \Delta R) + \hat{\mu}^{-2}(\Delta\bar{x}_n^2 \Delta M)\hat{a}\right),$$

21. $$\hat{\mu}_s^{-2}(\Delta\bar{x}_n^2 \Delta a_i \Delta R_i) = \hat{\mu}_s^{-1}(\Delta\bar{x}_n^2)\hat{\mu}_s^{-1}(\Delta a_i \Delta R_i) + 2\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta a_i)\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta R_i),$$

22. $$\hat{\mu}^{-2}(\Delta\bar{x}_n^2 \Delta(\sigma^2)) = \hat{\mu}^{-2}(\Delta\bar{x}_n^2 \Delta R_0) + \hat{a}^T \hat{\mu}^{-2}(\Delta\bar{x}_n^2 \Delta R) + R^T \hat{\mu}^{-2}(\Delta\bar{x}_n^2 \Delta a) + \sum_{i=1}^{k} \hat{\mu}_s^{-2}(\Delta\bar{x}_n^2 \Delta a_i \Delta R_i),$$

23. $$\hat{\mu}_s^{-1}(\Delta a \Delta a_i) = -\sum_{l=0}^{k} W_l^{(1)} \hat{\mu}_s^{-1}(\Delta a_i \Delta R_l),$$

24. $$\hat{\mu}_s^{-1}\left(\Delta(1^T a)^2\right) = \sum_{i,j=1}^{k} \hat{\mu}_s^{-1}(\Delta a_i \Delta a_j),$$

25. $$\hat{\mu}_s^{-1}(\Delta(\sigma^2)\Delta(1^T a)) = \sum_{i=1}^{k}\left[\hat{\mu}_s^{-1}(\Delta a_i \Delta R_0) + \sum_{j=1}^{k} \hat{R}_j \hat{\mu}_s^{-1}(\Delta a_i \Delta a_j) + \sum_{j=1}^{k} \hat{a}_j \hat{\mu}_s^{-1}(\Delta a_i \Delta R_j)\right],$$

26. $$\hat{\mu}_s^{-2}\left(\Delta\bar{x}_n^2\Delta(\sigma^2)\Delta(1^Ta)\right) = \hat{\mu}_s^{-1}\left(\Delta\bar{x}_n^2\right)\hat{\mu}_s^{-1}\left(\Delta(\sigma^2)\Delta(1^Ta)\right) + 2\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n\Delta(1^Ta)\right)\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n\Delta(\sigma^2)\right)$$

27. $$\hat{\mu}_s^{-2}\left(\Delta\bar{x}_n^2\Delta(1^Ta)^2\right) = \hat{\mu}_s^{-1}\left(\Delta\bar{x}_n^2\right)\hat{\mu}_s^{-1}\left(\Delta(1^Ta)^2\right) + 2\left(\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n\Delta(1^Ta)\right)\right)^2$$

28. $$\hat{\mu}^{-2}\left(\Delta\bar{x}_n^2\Delta v\right) = \frac{\hat{\mu}^{-2}\left(\Delta\bar{x}_n^2\Delta(\sigma^2)\right)}{\hat{A}(1)^2} - \frac{2\hat{\sigma}^2\hat{\mu}^{-2}\left(\Delta\bar{x}_n^2\Delta(1^Ta)\right)}{\hat{A}(1)^3} - \frac{2\hat{\mu}_s^{-2}\left(\Delta\bar{x}_n^2\Delta(\sigma^2)\Delta(1^Ta)\right)}{\hat{A}(1)^3} + 3\hat{\sigma}^2\frac{\hat{\mu}_s^{-2}\left(\Delta\bar{x}_n^2\Delta(1^Ta)^2\right)}{\hat{A}(1)^4}.$$

Next we will compute $\hat{\mu}_s^{-2}\left(\Delta\bar{x}_n^2\Delta v^2\right)$.

29. $$\hat{\mu}_s^{-1}\left(\Delta(\sigma^2)^2\right) = \sum_{i,j=0}^{k}\Big(\hat{a}_i\hat{a}_j\hat{\mu}_s^{-1}(\Delta R_i\Delta R_j) + \hat{R}_i\hat{R}_j\hat{\mu}_s^{-1}(\Delta a_i\Delta a_j) + 2\hat{a}_i\hat{R}_j\hat{\mu}_s^{-1}(\Delta a_j\Delta R_i)\Big),$$

where $\hat{\mu}_s^{-1}(\Delta a_0\Delta(\cdot)) \triangleq 0$ because $a_0 = 1$ is known.

30. $$\hat{\mu}_s^{-1}\left(\Delta v^2\right) = \frac{\hat{\mu}_s^{-1}\left(\Delta(\sigma^2)^2\right)}{\hat{A}(1)^4} + \frac{4\hat{\sigma}^4\hat{\mu}_s^{-1}\left(\Delta(1^Ta)^2\right)}{\hat{A}(1)^6}$$

$\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta M_2) \leftarrow \hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta a_k) I$

$\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta a_0) \triangleq 0$

For $l = k-1$ down to 0,

$$\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta M_2) \leftarrow \hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta F) M_2 + F\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta M_2) + \hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta a_l) I$$

$$M_2 \leftarrow FM_2 + \hat{a}_l I.$$

36. $$\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n \Delta(M_2^{-1})\right) = -M_2^{-1}\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta M_2) M_2^{-1}$$

37. $$M_3 = \begin{bmatrix} a_0 & & 0 \\ \vdots & \ddots & \\ \hat{a}_{k-1} & \dots & a_0 \end{bmatrix}$$

The coefficient of covariance of the point estimate with M_3, namely $\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta M_3)$, is

38. $$\begin{bmatrix} \hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta a_0) & & 0 & \\ \vdots & & \ddots & \\ \dots & \hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta a_{k-1}) & \dots & \hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta a_0) \end{bmatrix}.$$

The diagonal elements in the matrix above are all zero, but they are written as they are to indicate the structure of the matrix.

39. $$\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n \Delta(M_3^{-1})\right) = -M_3^{-1}\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta M_3) M_3^{-1}$$

40. $$\begin{aligned} S &= [S_0, \dots, S_{k-1}]^T \\ &= \left[\sum y_i^2, \dots, \sum y_i y_{k-i-1}\right]^T \end{aligned}$$

41. $$\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta S) = \hat{\mu}_s^{-1}\left(\Delta\bar{x}_n \Delta(M_2^{-1})\right) M_3^{-1} e_1 + M_2^{-1}\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n \Delta(M_3^{-1})\right) e_1.$$

$$- \frac{4\hat{\sigma}^2 \hat{\mu}_s^{-1}\big(\Delta(\sigma^2)\Delta(1^T a)\big)}{\hat{A}(1)^5},$$

31.
$$\hat{\mu}_s^{-2}\big(\Delta\bar{x}_n^2 \Delta v^2\big) = \hat{\mu}_s^{-1}\big(\Delta\bar{x}_n^2\big)\hat{\mu}_s^{-1}\big(\Delta v^2\big) + 2\left(\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta v)\right)^2 .$$

This enables us to compute

32.
$$\hat{E}t_n^2 - 1 = \frac{\hat{\mu}^{-2}\big(\Delta\bar{x}_n^2\big)}{n\hat{\mu}_s^{-1}\big(\Delta\bar{x}_n^2\big)} - \frac{\hat{\mu}^{-2}\big(\Delta\bar{x}_n^2 \Delta v\big)}{n\left(\hat{\mu}_s^{-1}\big(\Delta\bar{x}_n^2\big)\right)^2} + \frac{\hat{\mu}_s^{-2}\big(\Delta\bar{x}_n^2 \Delta v^2\big)}{n\left(\hat{\mu}_s^{-1}\big(\Delta\bar{x}_n^2\big)\right)^3}.$$

33.
$$\hat{\mu}_s^{-1}\big(\Delta\bar{x}_n \Delta(n^{-1}\textstyle\sum \epsilon_i^3)\big) = \mathrm{SUM}_{01}\Big(0, \ldots, 0, \hat{E}\epsilon^4\Big),$$

34.
$$\hat{\mu}_s^{-1}\Big(\Delta\bar{x}_n \Delta(\hat{E}\epsilon^3)\Big) = \hat{\mu}_s^{-1}\big(\Delta\bar{x}_n \Delta(n^{-1}\textstyle\sum \epsilon_i)\big) - 3\hat{\sigma}^2 \hat{A}(1)\hat{\mu}_s^{-1}\big(\Delta\bar{x}_n^2\big).$$

The next series of computations will give us $M_2 \triangleq (I + \hat{a}_1 F + \cdots + \hat{a}_k F^k)$ and $\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta M_2)$. The matrix $\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta F)$ below is a $k \times k$ matrix identically zero except in the last row:

$$\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta F) = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ -\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta a_k) & \ldots & -\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta a_1) \end{bmatrix}.$$

35. $M_2 \leftarrow \hat{a}_k I$

In the following calculations we will compute coefficients of covariances between the point estimate and estimated coefficients from the first order correction.

42.
$$\left[\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n\Delta\mu_s^{-1}(\Delta\bar{x}_n\Delta R_i)\right)\right]_{i=0}^{k-1} = \frac{S\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n\Delta(E\epsilon^3)\right)}{\hat{A}(1)} - \frac{S\hat{E}\epsilon^3\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n\Delta(1^T a)\right)}{\hat{A}(1)^2} + \frac{\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta S)\hat{E}\epsilon^3}{\hat{A}(1)},$$

43.
$$\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n\Delta\mu_s^{-1}(\Delta\bar{x}_n\Delta R_k)\right) = \sum_{l=1}^{k}\Big(-\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta a_l)\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta R_{k-l}) - \hat{a}_l\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n\Delta(\mu_s^{-1}(\Delta\bar{x}_n\Delta R_{k-l}))\right)\Big).$$

Define $\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n\Delta\mu_s^{-1}(\Delta\bar{x}_n\Delta R)\right)$ in the obvious way, that is by $\left[\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n\Delta\mu_s^{-1}(\Delta\bar{x}_n\Delta R_i)\right)\right]_{i=1}^{k}$, and similarly define the coefficient $\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n\Delta\mu_s^{-1}(\Delta\bar{x}_n\Delta\dot{M})\right)$.

44.
$$\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n\Delta(M^{-1})\right) = -\hat{M}^{-1}\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta M)\hat{M}^{-1}.$$

45.
$$\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n\Delta\mu_s^{-1}(\Delta\bar{x}_n\Delta a)\right) = -\hat{\mu}_s^{-1}\left(\Delta\bar{x}_n\Delta M^{-1}\right)\Big[\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta R)$$

$$
\begin{aligned}
&+ \hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta M)\hat{a}\Big] \\
&- \hat{M}^{-1}\Big[\hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta \mu_s^{-1}(\Delta \bar{x}_n \Delta R)\big) \\
&+ \hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta \mu_s^{-1}(\Delta \bar{x}_n \Delta M)\big)\hat{a} \\
&+ \hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta M)\hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta a)\Big]
\end{aligned}
$$

46.
$$
\begin{aligned}
&\hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta \mu_s^{-1}(\Delta \bar{x}_n \Delta(\sigma^2))\big) = \\
&\hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta \mu_s^{-1}(\Delta \bar{x}_n \Delta R_0)\big) \\
&\quad + \hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta R^T\big)\hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta a) \\
&\quad + \hat{R}^T \hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta \mu_s^{-1}(\Delta \bar{x}_n \Delta a)\big) \\
&\quad + \hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta a^T\big)\hat{\mu}_s^{-1}(\Delta \bar{x}_n \Delta R) \\
&\quad + \hat{a}^T \hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta \mu_s^{-1}(\Delta \bar{x}_n \Delta R)\big),
\end{aligned}
$$

47.
$$
\begin{aligned}
&\hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta \mu_s^{-1}(\Delta \bar{x}_n \Delta v)\big) = \\
&\frac{\hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta \mu_s^{-1}(\Delta \bar{x}_n \Delta(\sigma^2))\big)}{\hat{A}(1)^2} \\
&\quad - \frac{4\hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta(\sigma^2)\big)\hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta(1^T a)\big)}{\hat{A}(1)^3} \\
&\quad + \frac{6\hat{\sigma}^2\left(\hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta(1^T a)\big)\right)^2}{\hat{A}(1)^4} \\
&\quad - \frac{2\hat{\sigma}^2 \hat{\mu}_s^{-1}\big(\Delta \bar{x}_n \Delta \mu_s^{-1}(\Delta \bar{x}_n \Delta(1^T a))\big)}{\hat{A}(1)^3}.
\end{aligned}
$$

48. $$w^T = \frac{[1 + \hat{a}_1 + \cdots + \hat{a}_{k-1}, \ldots, 1 + \hat{a}_1, 1]}{\hat{A}(1)},$$

49. $$\begin{aligned}\hat{\mu}_s^{-1}\big(\Delta\bar{x}_n \Delta w^T\big) &= [1 + \hat{a}_1 + \cdots + \hat{a}_{k-1}, \ldots, 1 + \hat{a}_1, 1] \\ &\quad \times \frac{(-1)\hat{\mu}_s^{-1}\big(\Delta\bar{x}_n \Delta(1^T a)\big)}{\hat{A}(1)^2} \\ &\quad + \frac{1}{\hat{A}(1)}\Bigg[\sum_{i=1}^{k-1} \hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta a_i), \\ &\qquad \ldots, \sum_{i=1}^{1} \hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta a_1), 0\Bigg]\end{aligned}$$

50. $$\begin{aligned}&\hat{\mu}_s^{-1}\big(\Delta\bar{x}_n \Delta\hat{\mu}^{-1}(\Delta\bar{x}_n)\big) = \\ &\quad - \mathrm{SUM}_{01}(1, \ldots, 1) \cdot \hat{\mu}_s^{-1}\big(\Delta\bar{x}_n^2\big) \\ &\quad + \hat{\mu}_s^{-1}\big(\Delta\bar{x}_n \Delta w^T\big) \cdot [x_1 - \bar{x}_n, \ldots, x_k - \bar{x}_n]^T.\end{aligned}$$

51. $$\begin{aligned}\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta\alpha_1) &= \hat{\mu}_s^{-1}\big(\Delta\bar{x}_n \Delta\mu^{-1}(\Delta\bar{x}_n)\big) v_n^{-1/2} \\ &\quad - \frac{1}{2} v_n^{-3/2} \hat{\mu}^{-1}(\Delta\bar{x}_n) \hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta v) \\ &\quad - \frac{1}{2} v_n^{-3/2} \hat{\mu}_s^{-1}\big(\Delta\bar{x}_n \Delta\mu_s^{-1}(\Delta\bar{x}_n \Delta v)\big) \\ &\quad + \frac{3}{4} v_n^{-5/2} \big(\hat{\mu}_s^{-1}(\Delta\bar{x}_n \Delta v)\big)^2.\end{aligned}$$

52. $$\begin{aligned}\hat{\mu}_s^{-1}\big(\Delta\bar{x}_n \Delta\mu_s^{-2}(\Delta\bar{x}_n^3)\big) &= \frac{\hat{\mu}_s^{-1}\Big(\Delta\bar{x}_n \Delta(\hat{E}\epsilon^3)\Big)}{\hat{A}(1)^3} \\ &\quad - \frac{3\hat{E}\epsilon^3 \hat{\mu}_s^{-1}\big(\Delta\bar{x}_n \Delta(1^T a)\big)}{\hat{A}(1)^4}.\end{aligned}$$

53. $$\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta\beta_1) = \hat{\mu}_s^{-1}\big(\Delta\bar{x}_n\Delta\mu_s^{-2}(\Delta\bar{x}_n^3)\big)v_n^{-3/2}$$
$$- \frac{3}{2}v_n^{-5/2}\hat{\mu}_s^{-2}(\Delta\bar{x}_n^3)\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta v)$$
$$+ 3\hat{\mu}_s^{-1}\big(\Delta\bar{x}_n\Delta\mu^{-1}(\Delta\bar{x}_n)\big)v_n^{-1/2}$$
$$- \frac{3}{2}v_n^{-3/2}\hat{\mu}^{-1}(\Delta\bar{x}_n)\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta v)$$
$$- \frac{9}{2}\hat{\mu}_s^{-1}\big(\Delta\bar{x}_n\Delta\mu_s^{-1}(\Delta\bar{x}_n\Delta v)\big)v_n^{-3/2}$$
$$+ \frac{27}{4}v_n^{-5/2}\left(\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta v)\right)^2.$$

54. $$\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta\theta) = -\frac{3}{2}\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta\alpha_1) + \frac{1}{6}\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta\beta_1)$$

55. $$\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta\rho) = \frac{1}{2}\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta\alpha_1) - \frac{1}{6}\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta\beta_1)$$

56. $$\hat{\mu}_s^{-1/2}(\Delta t_n{}^3\Delta\rho) = 3\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta\rho)v_n^{-1/2}$$

57. $$\hat{\mu}_s^{-1/2}(\Delta t_n\Delta\theta) = \hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta\theta)v_n^{-1/2}.$$

This enables us at last to compute the estimated second order cumulant of T_1, namely $\hat{\kappa}_2(T_1)$:

58. $$\hat{\kappa}_2(T_1) - 1 = \hat{E}t_n^2 - 1 + \frac{2\hat{\mu}_s^{-1/2}(\Delta t_n{}^3\Delta\rho)}{n}$$
$$+ \frac{2\hat{\mu}_s^{-1/2}(\Delta t_n\Delta\theta)}{n} + \frac{\hat{\theta}^2}{n} + \frac{3\hat{\rho}^2}{n}$$
$$+ \frac{2\hat{\theta}\hat{\alpha}_1}{n} + \frac{2\hat{\rho}\hat{\beta}_1}{n} + \frac{2\hat{\theta}\hat{\rho}}{n}.$$

Now we turn to the fourth order cumulant of T_1.

59. $$\hat{\kappa}_4(\epsilon) = \hat{E}\epsilon^4 - 3\hat{\sigma}^4$$

60. $$\hat{\mu}_s^{-3}(\Delta\bar{x}_n^4) = \frac{\hat{\kappa}_4(\epsilon)}{\hat{A}(1)^4} + 6\hat{\mu}_s^{-1}(\Delta\bar{x}_n^2)\hat{\mu}_s^{-2}(\Delta\bar{x}_n^2)$$

61. $$\hat{\mu}^{-3}(\Delta\bar{x}_n^4) = \hat{\mu}_s^{-3}(\Delta\bar{x}_n^4) + 4\hat{\mu}^{-1}(\Delta\bar{x}_n)\hat{\mu}_s^{-2}(\Delta\bar{x}_n^3) + 6\hat{\mu}_s^{-1}(\Delta\bar{x}_n^2)\left(\hat{\mu}^{-2}(\Delta\bar{x}_n^2) - \hat{\mu}_s^{-2}(\Delta\bar{x}_n^2)\right).$$

By $\hat{\mu}^{-1}(\Delta R_j)$ we mean $\lim_{n\to\infty} n(\hat{R}_j - R_j)$. For $j = 0, \ldots, k$ compute

62. $$\hat{\mu}^{-1}(\Delta R_j) = \hat{\mu}^{-1}\left(\Delta R_j'\right) - jR_j - \hat{\mu}_s^{-1}(\Delta\bar{x}_n^2).$$

63. $$\hat{\mu}^{-1}(\Delta a) = -\hat{M}^{-1}\left(\hat{\mu}^{-1}(\Delta R) + \hat{\mu}^{-1}(\Delta M)a\right) + \sum_{i,j=0}^{k}\hat{\mu}_s^{-1}(\Delta R_i\Delta R_j)W_{i,j}^{(2)},$$

64. $$\hat{\mu}^{-1}\left(\Delta(\sigma^2)\right) = \hat{\mu}^{-1}(\Delta R_0) + a^T\hat{\mu}^{-1}(\Delta R) + R^T\hat{\mu}^{-1}(\Delta a) + \sum_{i=1}^{k}\hat{\mu}_s^{-1}(\Delta a_i\Delta R_i),$$

65. $$\hat{\mu}^{-1}(\Delta v) = \frac{\hat{\mu}^{-1}(\Delta(\sigma^2))}{\hat{A}(1)^2} - \frac{2\hat{\sigma}^2 1^T\hat{\mu}^{-1}(\Delta a)}{\hat{A}(1)^3} - \frac{2\hat{\mu}_s^{-1}(\Delta(\sigma^2)\Delta(1^T a))}{\hat{A}(1)^3} + 3\hat{\sigma}^2\frac{\hat{\mu}_s^{-1}\left(\Delta(1^T a)^2\right)}{\hat{A}(1)^4}.$$

66. $$\hat{\mu}^{-3}(\Delta\bar{x}_n^4\Delta v) = 4\hat{\mu}_s^{-2}(\Delta\bar{x}_n^3)\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta v) + 6\hat{\mu}^{-2}(\Delta\bar{x}_n^2\Delta v)v_n - 3v_n^2\hat{\mu}^{-1}(\Delta v),$$

67. $$\hat{\mu}_s^{-3}(\Delta\bar{x}_n^4\Delta v^2) = 3\left(\hat{\mu}_s^{-1}(\Delta\bar{x}_n^2)\right)^2\hat{\mu}_s^{-1}(\Delta v^2) + 12\hat{\mu}_s^{-1}(\Delta\bar{x}_n^2)\left(\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta v)\right)^2,$$

68. $$\hat{E}t_n^4 - 3 = \frac{1}{n}\left[\frac{\hat{\mu}^{-3}(\Delta\bar{x}_n^4)}{v_n^2} - \frac{2\hat{\mu}^{-3}(\Delta\bar{x}_n^4\Delta v)}{v_n^3} + \frac{3\hat{\mu}_s^{-3}(\Delta\bar{x}_n^4\Delta v^2)}{v_n^4}\right]$$

69. $$\hat{\mu}^{-1/2}(\Delta t_n{}^3\Delta\theta) = \hat{\mu}_s^{-2}(\Delta\bar{x}_n^3\Delta\theta)v_n^{-3/2} = 3\hat{\mu}_s^{-1}(\Delta\bar{x}_n\Delta\theta)v_n^{-1/2}.$$

In the following, note that $\hat{\mu}^{-1/2}(\Delta t_n{}^3)$ is the same (by definition) as $\hat{\beta}_1$, and $\hat{\mu}^{-1/2}(\Delta t_n)$ is the same as $\hat{\alpha}_1$.

70. $$\hat{\mu}^{-1/2}(\Delta t_n{}^5) = 10\hat{\mu}^{-1/2}(\Delta t_n{}^3) - 15\hat{\mu}^{-1/2}(\Delta t_n),$$

71. $$\hat{\mu}^{-1/2}(\Delta t_n{}^5\Delta\rho) = 15\hat{\mu}_s^{-1/2}(\Delta x_n\Delta\rho)v_n^{-1/2},$$

72. $$\hat{\mu}^0(\Delta t_n{}^6) = 15.$$

73. $$\hat{E}T_1^4 - 3 = \hat{E}t_n^4 - 3 + \frac{1}{n}\Big[4\hat{\theta}\hat{\mu}^{-1/2}(\Delta t_n{}^3) + 4\hat{\mu}^{-1/2}(\Delta t_n{}^3\Delta\theta) + 4\hat{\rho}\hat{\mu}^{-1/2}(\Delta t_n{}^5) + 4\hat{\mu}^{-1/2}(\Delta t_n{}^5\Delta\rho) + 6\hat{\theta}^2 + 90\hat{\rho}^2 + 36\hat{\theta}\hat{\rho}\Big].$$

And therefore the fourth order cumulant estimate for T_1 is

74. $$\hat{\kappa}_4(T_1) = (\hat{E}T_1^4 - 3) - 6(\hat{\kappa}_2(T_1) - 1).$$

From this the second order correction T_2 is computed as

75. $$T_2 = T_1 + \frac{t_n\hat{\nu}_n}{n} + \frac{t_n^3\hat{\omega}_n}{n}$$

where

$$\frac{\hat{\nu}_n}{n} \triangleq -\frac{1}{2}(\hat{\kappa}_2(T_1) - 1) + \frac{1}{8}\hat{\kappa}_4(T_1)$$

and

$$\frac{\hat{\omega}_n}{n} \triangleq -\frac{1}{24}\hat{\kappa}_4(T_1).$$

The quantiles t_p for t_n may be estimated from corresponding standard normal quantiles z_p by

76. $$t_p = z_p - \frac{1}{\sqrt{n}}\left(\hat{\theta} + \hat{\rho}z_p^2\right) + \frac{1}{n}\left(-\hat{\nu}z_p + 2\hat{\theta}\hat{\rho}z_p - \hat{\omega}z_p^3 + 2\hat{\rho}^2z_p^3\right).$$

A $(1-p)$ confidence interval for $\mu = E_s x$ is then

77. $$(\bar{x}_n - n^{-1/2}v_n^{1/2}t_{1-p/2},\ \bar{x}_n + n^{-1/2}v_n^{1/2}t_{p/2}).$$

4.9 A note on validation of algebra. If for a given autoregressive model one inputs the true covariances and moments of the residuals into the algorithms for the first and second order pivots, one can find the true values of all the quantities estimated in this chapter, provided the calculations here are correct. Even for the zero order pivot, using true covariances will yield the true asymptotic variance constant v. This was done for an AR3 model, and the theoretical results thus obtained were compared with simulated values. Simulated values are of course the correct ones, except for the error of estimation. In this experiment, there were 1,000 data points per replication and 10,000 replications. Here are a few results from the second order correction.

Quantity	*Theoretical*	*Simulated*
$\mu^{-2}(\Delta x^2 \Delta R_0)$	$22.23 \cdot 10^1$	$22.35 \cdot 10^1$
$\mu^{-2}(\Delta R_3{}^3)$	83.62	81.06
$\mu_s^{-1}(\Delta a_3 \Delta R_3)$	$-20.45 \cdot 10^{-1}$	$-19.76 \cdot 10^{-1}$
$\mu^{-2}(\Delta x^2 \Delta a_1)$	-25.94	-23.13
$\mu_s^{-1}(\Delta x \Delta \mu_s^{-1}(\Delta x \Delta R_0))$	-28.65	-29.76
$\mu_s^{-1}(\Delta x \Delta \mu_s^{-1}(\Delta x \Delta(\sigma^2)))$	-24.0	-25.02
$\mu_s^{-3}(\Delta x^4 \Delta v^2)$	$32.56 \cdot 10^4$	$32.86 \cdot 10^4$
$\mu^{-3}(\Delta x^4 \Delta v)$	$36.83 \cdot 10^3$	$48.90 \cdot 10^3$
$ET_1^4 - 3$	$15.90 \cdot 10^{-2}$	$-7.14 \cdot 10^{-2}$

The above figures are typical. Note that in the last two rows above, the agreement may not be as good as one might expect, but the numbers are in fact acceptable. $ET_1^4 - 3$ is one of several second order quantities with an extremely high coefficient of variation. Because of this, it is not possible to obtain a good estimate given current computing constraints. Estimates of odd order moments like $\mu^{-3}(\Delta x^4 \Delta v)$ also tend to have a high coefficient of variation. In summary, almost all of the estimates obtained confirm the algebra of this chapter, and none of them arouses suspicion.

5

Numerical results

Introduction

The usual test statistic t_n is asymptotically normal, as are the corrected statistics T_1 and T_2. The distributions of the latter two statistics, however, converge more quickly to the standard normal distribution. (See for example ABRAMOVITCH AND SINGH (1985) and the references cited there).

To understand the data to follow, it may be helpful to give an informal review of how we form the usual, "zero order" confidence interval based on the standard test statistic t_n, and how we form the first and second order corrected confidence intervals.

The statistic t_n converges weakly to a standard normal random variable. If z_p denotes the p-quantile of the standard normal, the approximation

$$P\{z_{.05} \leq t_n \leq z_{.95}\} \approx .9$$

becomes more accurate as n increases ("$\approx$" means approximately equal). In fact, the error is $O(n^{-1/2})$. The event that the true asymptotic mean μ lies in the nominal 90 percent confidence interval constructed from t_n is precisely the event

$$\{z_{.05} \leq t_n \leq z_{.95}\},$$

while, for example, the event that μ lies above the upper 90 percent confidence interval bound is the same as the event

$$\{t_n \leq z_{.05}\}.$$

This is evident by rewriting t_n as $\sqrt{n}(\bar{x}_n - \mu)/\hat{\sigma}_n$.

T_1 is a quadratic polynomial in t_n, and T_2 is a cubic polynomial in t_n, which we might indicate by writing $T_1(t_n)$ and $T_2(t_n)$. For values of t_n of interest and for sufficiently large n, these polynomials are approximately the identity function:

$$T_1(t_n) = t_n + O_p(n^{-1/2})$$
$$T_2(t_n) = T_1(t_n) + O_p(n^{-1}).$$

These polynomial transformations may be regarded as transforming the distribution of t_n into a (more) normal distribution. Equivalently, these transformations map quantiles of the distribution of t_n into corresponding approximate normal quantiles. Not that t_n, T_1, or T_2 are actually normal, but each of the following statements is more accurate than its predecessor:

$$P\{z_{.05} \leq t_n \leq z_{.95}\} \approx .9,$$
$$P\{z_{.05} \leq T_1 \leq z_{.95}\} \approx .9,$$
$$P\{z_{.05} \leq T_2 \leq z_{.95}\} \approx .9.$$

In fact, the errors are $O(n^{-1/2})$, $o(n^{-1/2})$, and $o(n^{-1})$, respectively.

The inverted transformations may be written as $T_1^{-1}(z_p)$ and $T_2^{-1}(z_p)$. Instead of converting t_n-quantiles to approximate normal quantiles, we are doing the reverse: converting normal quantiles to approximate t_n-quantiles. The inverses of the quadratic $T_1(\cdot)$ and of the cubic $T_2(\cdot)$ are not polynomials, but we have shown how to approximate these inverses with a polynomial up to the desired order of accuracy. These inverted polynomials are then $T_2^{-1}(\cdot)$ and $T_1^{-1}(\cdot)$. They, too, are nearly the identity for fixed values of their arguments in the range of interest:

$$T_1^{-1}(z) = z + O_p(n^{-1/2})$$
$$T_2^{-1}(z) = T_1^{-1}(z) + O_p(n^{-1}).$$

We could base modified confidence intervals on the statements

$$P\{z_{.05} \leq T_1 \leq z_{.95}\} \approx .9$$
$$P\{z_{.05} \leq T_2 \leq z_{.95}\} \approx .9,$$

but conversion of these statements into statements about $(\bar{x}_n - \mu)$ requires solving a quadratic or cubic equation. It is therefore more convenient to base intervals on

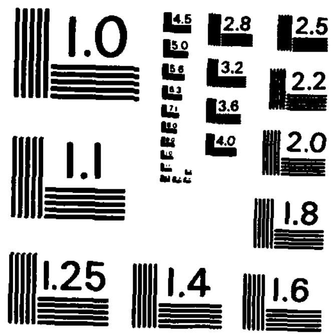
$$P\{T_1^{-1}(z_{.05}) \leq t_n \leq T_1^{-1}(z_{.95})\} \approx .9$$
$$P\{T_2^{-1}(z_{.05}) \leq t_n \leq T_2^{-1}(z_{.95})\} \approx .9.$$

The event that the true mean μ lies above the lower 90 percent confidence interval bound obtained from the first order corrected interval, for example, is the same as the event that $t_n \leq T_1^{-1}(z_{.05})$.

AD-A160 055 MODIFIED CONFIDENCE INTERVALS FOR THE MEAN OF AN AUTOREGRESSIVE PROCESS(U) STANFORD UNIV CA DEPT OF OPERATIONS RESEARCH B D TITUS AUG 85 TR-9 2/2

UNCLASSIFIED ARO-20927.9-MA DAAG29-84-K-0030 F/G 12/1 NL

END
FILMED
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

In this chapter we will compare the distribution of the usual test statistic t_n with the standard normal distribution and with the estimated distribution of t_n obtained from the inverted versions of T_1 and T_2, which we denote by T_1^{-1} and T_2^{-1}. As we have said, in practice one does not use T_1 or T_2 directly to form a confidence interval because this would require solving a cubic or quadratic equation. Instead, we use the inverted corrections discussed above and given in equations (4.7.16) or (4.8.76). However, the comparison of T_1 or T_2 with the standard normal is qualitatively similar to the comparison of the inverted statistics with t_n.

It would be possible for us just to report the true coverage probability for various intervals. In other words, with what frequency does the true asymptotic mean fall above a 90 percent confidence interval, or below it, or in the upper half, or in the lower half? There is, we feel, a more informative way to compare the various pivots using what we will call *p-p plots* which are a variation on *q-q plots*.

Suppose we wish to compare two distributions, F_1 and F_2. A *q-q* plot plots the two quantiles corresponding to a given probability. Thus the plot includes points of the form $(F_1^{-1}(p), F_2^{-1}(p))$. If F_1 is standard normal, and F_2 is an empirical distribution, this is well known as a useful way to test normality. It has the virtue of testing the tails (which are typically of most interest to us) and also the virtue that if F_2 is non-standard normal the plot will still be a *straight line*.

For our purposes, the *p-p* plot is better. Here we plot points

$(F_1(q), F_2(q))$. Both axes of these plots consist of the interval $[0, 1]$. Suppose F_2 is standard normal and F_1 is the distribution of t_n, and that the point (p_1, p_2) appears on the graph. This means that the p_1 quantile of t_n is the same number as the p_2 quantile of the standard normal, which in turn means that the true probability that t_n is less than the p_2-normal quantile is p_1. The true coverage probability for a 90 percent confidence interval based on t_n is then $p_{\mathrm{high}} - p_{\mathrm{low}}$ where the points $(p_{\mathrm{high}}, 0.95)$ and $(p_{\mathrm{low}}, 0.05)$ appear on the plot.

Referring to the first graph for Model 1, four increments in normal probability subdivide $[0, 1]$ on the y-axis ("Normal p" axis), namely 0.05, 0.45, 0.45, and 0.05. The corresponding increments are indicated along the t_n p axis, which in this case are 0.18, 0.47, 0.32 and 0.03. This means, for example, that the probability that t_n is less than or equal to the 0.05 normal quantile is actually 0.18. Furthermore, the event that t_n is less than the 0.05 normal quantile is the same as the event that the true mean μ lies *above* the upper 90 percent confidence limit. For reference, the dots on the graph indicate the points $(.1, .1), \ldots, (1, 1)$.

The results of testing the inverted corrections are displayed in a similar way. The inverted first order statistic, T_1^{-1}, gives for each replication an estimate of any quantile of t_n. This estimate of the pth quantile is $T_1^{-1}(z_p)$, and it changes with each replication. Referring to the second row of the tables for Model 1, the t_n p coordinate corresponding to the 0.95 T_1^{-1} p coordinate gives the frequency with which t_n is less than or equal to its estimated 0.95 quantile obtained from the first order correction, namely $T_1^{-1}(z_{.95})$. Again, the event that t_n is less than or equal to this estimate is the same as the event that the true mean μ lies above the lower 90 percent confidence interval bound as given by the first order corrected method. In this way, we see that for $n = 200$ data points, the probabilities that a nominal 90 percent confidence interval covers the true mean are 0.79, 0.81, and 0.87 for the zero, first, and second order methods, respectively.

Models tested

We will test the following four models.

(1) Autoregressive model $(x_i - 10) - 0.5(x_{i-1} - 10) - 0.3(x_{i-2} - 10) - 0.1(x_{i-3} - 10) = \epsilon_i - 1$, where the ϵ_i's are exponential with mean 1. The three initial values, x_1, x_2, and x_3, are set to seven, which is three less than the true mean. Results are shown for $n = 200$ and $n = 400$ data points.

(2) The same model as (1), but with geometric residuals: $P\{\epsilon_i = j\} = 2^{-j-1}$, for $j \geq 0$.

(3) The waiting time process in the M/M/1 queue, $\{W_i : i \geq 1\}$, with traffic intensity $\rho = 0.5$. We model this as an AR5 process and set the initial value to 0. Results are shown for $n = 1{,}000$ data points.

(4) A Markov chain on the nonnegative integers with transition probabilities $p_{i,i+1} = 1 - p_{i,i-1} = 1/3$ for $i \geq 1$. This is a discrete analog of the queue length process of the third model,

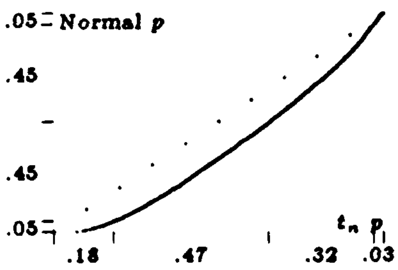
which we also model as AR5. The initial value is set to zero. Results are shown for $n = 1{,}000$ data points.

Note that the third model violates the independence assumption on the prediction errors and also the assumption of finite autoregressive order. We therefore would not necessarily expect good results for this model. In the second and fourth models, either the prediction errors or the process itself has a lattice distribution, which again means we would not necessarily expect good results.

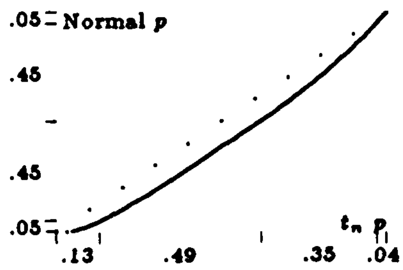
Data

We have chosen to do 10,000 replications, a number somewhat larger than those used in JOHNSON (1978), GLYNN (1982a), and JOW (1982). Suppose we want to estimate the probability that t_n is less than or equal to the 0.05 quantile of the standard normal distribution, $z_{.05}$. Our experiment is one of binomial trials with p approximately equal to 0.05. A 95 percent confidence interval for the required probability will then have halfwidth of 0.0043, or about half a percent. This is, we feel, a desirable level of accuracy for such an experiment. The halfwidths corresponding to $z_{.50}$ and $z_{.01}$ are about 1.0 percent and 0.2 percent.

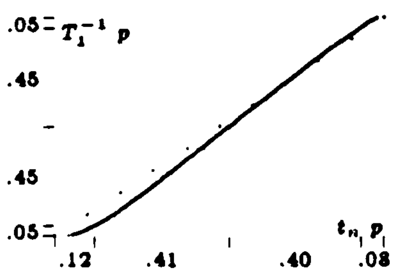
5.1 Model 1.



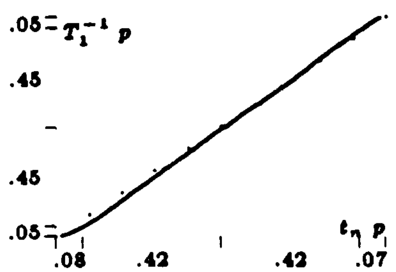
$n = 200$



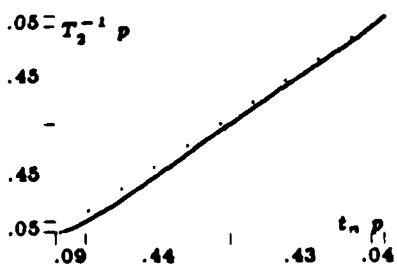
$n = 400$



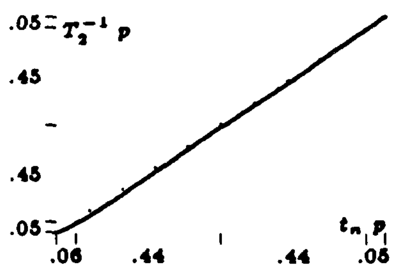
$n = 200$



$n = 400$

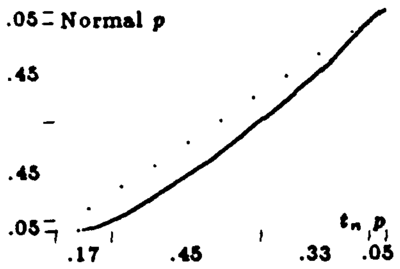


$n = 200$

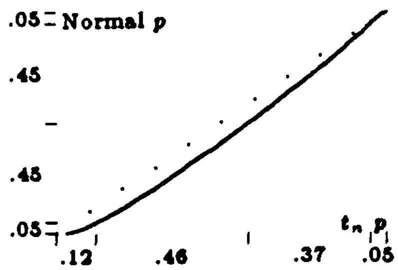


$n = 400$

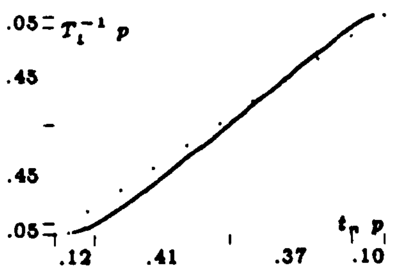
5.2 Model 2.



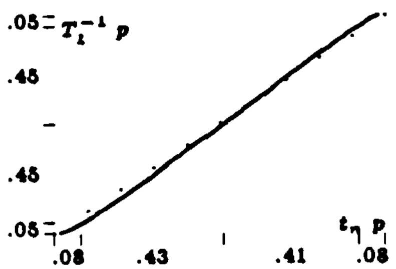
$n = 200$



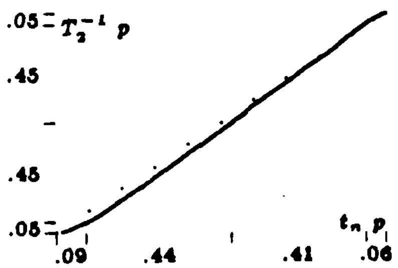
$n = 400$



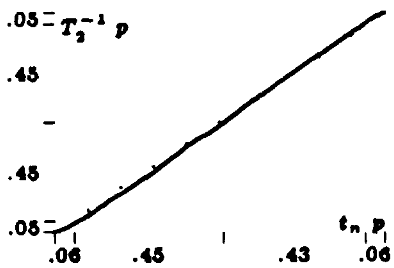
$n = 200$



$n = 400$

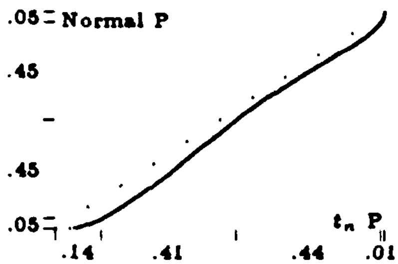


$n = 200$

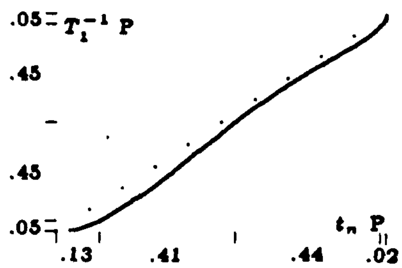


$n = 400$

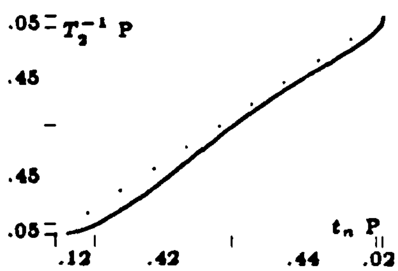
5.3 Model 3.



$n = 1000$

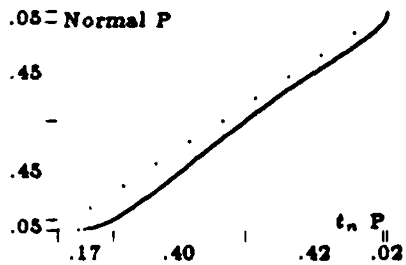


$n = 1000$

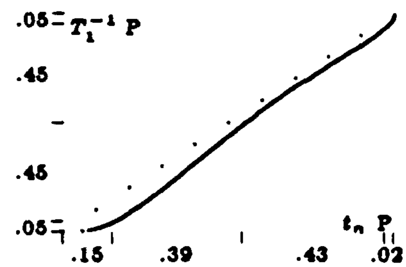


$n = 1000$

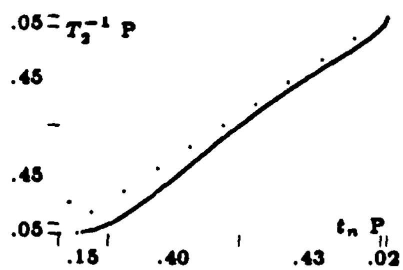
5.4 Model 4.



$n = 1000$



$n = 1000$



$n = 1000$

5.5 Discussion of data. The graphs indicate that the method performs very well for both of the true autoregressive models tested. This is to be expected, especially in the case of the first model which satisfies all of the sufficient criteria posed in Chapter 4. In the latter two models, we also see some improvement despite the dependence of residuals, but we would expect to see more improvement if this dependence were taken into account. Using the Fast Fourier Transform, it is possible to perform the first order correction in the case of dependent residuals in $O(n \log(n))$ time, but a second order correction would be more difficult. This is one possible area of future research. Autoregressive models have the virtue of tractability and of being a convenient way to take into account some kind of nonstationary behavior. But in many cases, some other method of modeling this nonstationarity may be more desirable, and the basic ideas and tools developed here can be used to obtain asymptotically more accurate confidence intervals in these other contexts.

References

ABRAMOVITCH, L. AND SINGH, K. (1985). Edgeworth corrected pivotal statistics and the bootstrap. ***Ann. Statist.* 13,** 116–132.

BHATTACHARYA, R. N. AND GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. ***Ann. Statist.* 6,** 434–451.

BHATTACHARYA, R. N. AND RANGA RAO, R. (1976). ***Normal Approximation and Asymptotic Expansions.*** Wiley, New York.

BIKJALIS, A. (1973). Asymptotic expansions for the densities and distributions of sums of independent and identically distributed random vectors, ***Selected Transl. in Math. Stat. and Probability* 13**, 213–234. The original article appeared in Russsian in 1968.

BILLINGSLEY, P. (1968). ***Convergence of Probability Measures.*** Wiley, New York.

BILLINGSLEY, P. (1979). ***Probability and Measure.*** Wiley, New York.

EDGEWORTH, F. Y. (1904). The law of error. ***Trans. Camb. Phil. Soc.* 20.**

EFRON, B. (1984a). Bootstrap confidence intervals for parametric problems. *Technical Report No. 215*, Department of Statistics, Stanford.

EFRON, B. (1984b). Better bootstrap confidence intervals. *Technical Report No. 14*, Department of Statistics, Laboratory for computational statistics, Stanford.

FISHMAN, G. S. (1978). ***Principles of Discrete Event Simulation.*** Wiley, New York.

FOX, B. AND GLYNN, P. (1983). Estimating time averages via randomly spaced observations. ***Technical report,*** Mathematics Research Center, University of Wisconsin..

GLYNN, P. (1982a). Asymptotic theory for nonparametric confidence intervals. ***Technical Report No. 63***, Department of Operations Research, Stanford.

GLYNN, P. (1982b). Regenerative aspects of the steady state simulation problem for Markov chains. ***Technical Report No. 61***, Department of Operations Research, Stanford.

GOLUB, G. H. AND VAN LOAN, C. F. (1983). ***Matrix Computations.*** Johns Hopkins, Baltimore.

GÖTZE, F. AND HIPP, C. (1983). Asymptotic expansions for sums of weakly dependent random vectors. ***Z. Wahrsch. Verw. Gebiete* 64,** 211–239.

HALL, P. (1983). Inverting an Edgeworth expansion. ***Ann. Statist.* 11,** 569–576.

HILL, G. AND DAVIS, A. (1968). Generalized asymptotic expansions of Cornish Fisher type. ***Ann. Math. Stat.* 39,** 1264–1273.

HOTELLING, A. AND FRANKEL, L. R. (1938). The transformation of statistics to simplify their distribution. ***Ann. Math. Stat.* 9,** 87–96.

JAMES, G. S. (1955). Cumulants of a transformed variate. *Biometrika* **42,** 529–531.

JAMES, G. S. (1958). On the moments and cumulants of systems of statistics. *Sankhyā Ser. A* **20,** 1–30.

JAMES, G. S. AND MAYNE, A. S. (1962). Cumulants of functions of random variables. *Sankhyā Ser. A* **24,** 47–54.

JOHNSON, N. J. (1978). Modified t-tests and confidence intervals for asymmetrical populations. *J. Amer. Statist. Assoc.* **73,** 536–544.

JOW, L. (1982). An autoregressive method for simulation output analysis. *Technical Report No. 21*, Department of Operations Research, Stanford.

KENDALL, M. AND STUART, A. (1977). *The Advanced Theory of Statistics, Vol. 1, 4th Edition.* Macmillan, New York.

LAW, A. AND KELTON, D. (1984). Confidence intervals for steady-state simulations: I. A survey of fixed sample size procedures. *Opns. Res.* **32,** 1221–1239.

PRIESTLEY, M. B. (1981). *The spectral analysis of time series, 2 vols..* Academic Press, London.

TANIGUCHI, M. (1984). Validity of Edgeworth expansions for statistics of time series. *J. Time Ser. Anal.* **15,** 37–51.

THOMSON, D. (1982). Spectrum Estimation and Harmonic Analysis. *Proc. IEEE* **70,** 1055–1096.

WITHERS, C. S. (1983). Expansions for the distribution and quantiles of a regular functional of the empirical distribution. *Ann. Statist.* **11,** 577–587.

END

FILMED

12-85

DTIC